Electromagnétisme et Relativité

J.M. Raimond
Laboratoire Kastler Brossel
Département de Physique de l'Ecole Normale Supérieure
jmr@physique.ens.fr

October 10, 2000

Table des Matières

1 Introduction générale **9**

I Mécanique analytique **15**

1 Formulation lagrangienne **19**
1.1 Description du système: coordonnées généralisées . . . 19
1.2 Principe de moindre action . . . 21
1.2.1 Enoncé . . . 21
1.2.2 Equations de Lagrange . . . 21
1.3 Expressions de la fonction de Lagrange . . . 23
1.3.1 Particule unique libre . . . 23
1.3.2 Système de particules interagissant par des forces dérivant d'un potentiel . . . 24
1.3.3 Système de particules soumises à des forces extérieures . . . 25
1.3.4 Lagrangien de particules chargées dans un champ . . . 27
1.4 Généralisations . . . 29
1.4.1 Forces ne dérivant pas d'une énergie potentielle . . . 30
1.4.2 Cas des liaisons non holonomes . . . 32
1.5 Lagrangien et lois de conservation . . . 35
1.5.1 Invariance par translation dans le temps: énergie . . . 36
1.5.2 Translation spatiale: conservation de l'impulsion . . . 37
1.5.3 Invariance par rotation: moment cinétique . . . 38
1.6 Action en fonction de la trajectoire . . . 39
1.6.1 Dépendance en position . . . 39
1.6.2 Dépendance en temps . . . 40
1.7 Deux théorèmes utiles . . . 41
1.7.1 Lois d'échelle . . . 41
1.7.2 Théorème du Viriel . . . 42

2 Formulation hamiltonienne **45**
2.1 Equations de Hamilton . . . 45
2.2 Crochets de Poisson . . . 48
2.3 Action et hamiltonien . . . 51
2.4 Transformations canoniques . . . 52
2.4.1 Principe . . . 52
2.4.2 Transformations canoniques et crochets de Poisson . . . 54
2.4.3 Exemples de transformations canoniques . . . 55
2.4.4 Transformations canoniques et espace des phases . . . 56
2.4.5 Transformation générée par l'action et théorème de Liouville . . . 58

Appendice 1 Modèle de Bohr **61**
1.1 Un peu d'histoire . . . 61
1.2 Modèle de Bohr . . . 63
1.3 Au delà du modèle de Bohr . . . 65

II Relativité restreinte **69**

1 Cinématique relativiste **73**
1.1 Rappels de relativité galiléenne . . . 73
1.1.1 Transformation de Galilée . . . 73
1.1.2 Les difficultés de la cinématique classique . . . 75
1.2 Principe de relativité . . . 76
1.2.1 Enoncé . . . 76
1.2.2 Deux expériences de pensée . . . 76
1.3 Evénements et intervalles . . . 79
1.3.1 Evénements . . . 79
1.3.2 Intervalle. Invariance de l'intervalle . . . 81
1.3.3 Conséquences de la conservation de l'intervalle. Causalité relativiste . . . 82
1.3.4 Temps propre . . . 82
1.4 Transformation de Lorentz . . . 85
1.4.1 Forme de la transformation de Lorentz . . . 85
1.5 Propriétés de la transformation de Lorentz . . . 89
1.5.1 Composition des transformations . . . 89
1.5.2 Vitesse, célérité et rapidité . . . 90
1.5.3 Géométrie de la transformation de Lorentz . . . 91
1.6 Conséquences de la transformation de Lorentz . . . 92
1.6.1 Retour sur nos expériences de pensée . . . 92
1.6.2 Intervalles et simultanéité . . . 92
1.6.3 Loi de composition des vitesses . . . 93
1.6.4 Contraction des longueurs . . . 94

2 Notations Quadridimensionnelles **97**
2.1 4–vecteur position d'un événement. . . . 97
2.1.1 Coordonnées contravariantes . . . 97
2.1.2 Coordonnées covariantes . . . 99
2.1.3 Coordonnées covariantes, contravariantes et dualité . . . 100
2.1.4 Changement de référentiel, changement de base . . . 100
2.2 Autres 4–vecteurs . . . 104
2.2.1 Définition . . . 104
2.2.2 4–vitesse, 4–impulsion, 4–accélération . . . 104
2.2.3 Densité de courant . . . 105
2.2.4 Vecteur d'onde . . . 106
2.3 Tenseurs . . . 107
2.3.1 Tenseurs contravariants . . . 108
2.3.2 Tenseurs covariants, tenseurs mixtes . . . 108
2.3.3 Vocabulaire et exemples . . . 109
2.4 Dérivation et analyse vectorielle . . . 110
2.4.1 Dérivation . . . 111
2.4.2 Analyse vectorielle . . . 111

2.4.3 Intégration . . . 112

3 Dynamique relativiste 115
3.1 Particule Libre . . . 115
3.2 Energie–impulsion . . . 118
3.3 Particule soumise à une force . . . 119
3.4 Conservation de l'énergie–impulsion. Application aux collisions . . . 120
3.4.1 Seuil de réaction . . . 121
3.4.2 Effet Compton . . . 122

4 Electromagnétisme relativiste 125
4.1 Particule libre dans un champ imposé . . . 126
4.1.1 Equations de Lagrange . . . 126
4.1.2 Tenseur champ électromagnétique . . . 128
4.1.3 Force de Lorentz . . . 129
4.1.4 Changements de référentiels pour le champ . . . 131
4.1.5 Invariants du champ électromagnétique . . . 132
4.1.6 Premier groupe d'équations de Maxwell . . . 133
4.2 Champ en fonction des sources . . . 134
4.2.1 Interaction champ–courant . . . 135
4.2.2 Lagrangien du champ . . . 136
4.2.3 Equations de Lagrange . . . 136
4.2.4 Equations de Maxwell . . . 137
4.3 Energie–impulsion du champ . . . 138
4.3.1 Tenseur énergie–impulsion . . . 138
4.3.2 Lois de conservation. Interprétation . . . 139
4.3.3 Applications . . . 142

III Propagation, diffraction 145

1 Potentiels retardés 149
1.1 Fonction de Green . . . 149
1.1.1 Position du problème . . . 149
1.1.2 Définition de la fonction de Green . . . 150
1.1.3 Approche qualitative . . . 151
1.2 Solution rigoureuse . . . 152
1.2.1 Fonction de Green . . . 152
1.2.2 Forme covariante . . . 155
1.2.3 Potentiels retardés . . . 155

2 Diffraction: approche théorique 157
2.1 Formule de Kirchhoff . . . 158
2.2 Principe de Huygens . . . 159
2.3 Diffraction de Fraunhofer . . . 163
2.3.1 Approximation paraxiale . . . 163
2.3.2 Différentes expressions de la figure de diffraction . . . 164
2.3.3 Généralisation à une transparence arbitraire . . . 165
2.3.4 Propriétés générales de la figure de diffraction . . . 165
2.3.5 Quelques exemples . . . 166

3 Applications de la diffraction **169**
3.1 Pouvoir de résolution des instruments d'optique . . . 169
3.2 Traitement optique du signal . . . 171
3.2.1 Filtrage spatial . . . 171
3.2.2 Méthode de Labeyrie . . . 173
3.3 Holographie . . . 176
3.3.1 Principe . . . 176
3.3.2 Applications . . . 178
3.3.3 Holographie et conjugaison de phase . . . 179
3.4 Limite de l'optique géométrique . . . 180

Appendice 1 Jauges **183**

Appendice 2 Espace réciproque **185**
2.1 Espace réciproque . . . 185
2.1.1 Transformation de Fourier . . . 185
2.1.2 Champs longitudinaux et transverses . . . 186
2.2 Variables normales . . . 187
2.2.1 Electromagnétisme dans l'espace libre . . . 188
2.2.2 Dynamique des variables normales . . . 188
2.2.3 Décomposition en ondes planes . . . 189

Appendice 3 Faisceaux gaussiens **191**
3.1 Construction d'un mode gaussien . . . 191
3.2 Propriétés essentielles . . . 194
3.3 Cavités laser . . . 196
3.4 Optique gaussienne . . . 199

Appendice 4 Approximation eikonale **201**
4.1 Eikonale . . . 201
4.2 Approximation de l'optique géométrique . . . 203
4.2.1 Equation de l'eikonale . . . 203
4.2.2 Notion de rayon lumineux . . . 204
4.3 Equation des rayons lumineux . . . 205
4.3.1 Rayons et eikonale . . . 205
4.3.2 Applications . . . 206
4.3.3 Principe de Fermat . . . 207

IV Sources du champ électromagnétique **209**

1 Rayonnement d'une particule en mouvement **213**
1.1 Potentiels de Liénard–Wiechert . . . 213
1.2 Champs rayonnés . . . 217
1.2.1 Dérivées utiles . . . 218
1.2.2 Champ électrique . . . 219
1.2.3 Champ magnétique . . . 220
1.2.4 Discussion physique . . . 220
1.3 Réaction de rayonnement . . . 223
1.3.1 Approche qualitative . . . 223
1.3.2 Force de réaction de rayonnement . . . 224

1.3.3 Application . . . 225
1.4 Rayonnement du dipôle . . . 227
1.4.1 Champs rayonnés . . . 227
1.4.2 Champs à une distance arbitraire . . . 232

2 Développement multipolaire du champ rayonné 237
2.1 Développement multipolaire du potentiel vecteur . . . 237
2.1.1 Notations . . . 237
2.1.2 Potentiel vecteur . . . 239
2.2 Termes multipolaires . . . 240
2.2.1 Ordre 0: Dipôle électrique . . . 240
2.2.2 Ordre 1: Dipôle magnétique, Quadripôle électrique . . . 241
2.3 Applications: quelques problèmes de rayonnement . . . 248
2.3.1 Rayonnement d'une charge oscillante . . . 248
2.3.2 Antennes . . . 250

3 Sources atomiques de rayonnement 255
3.1 Modèle de Thomson . . . 256
3.1.1 Modèle de l'électron élastiquement lié. . . . 256
3.1.2 Emission spontanée . . . 257
3.1.3 Diffusion du rayonnement. . . . 259
3.2 Modèle semi–classique . . . 264
3.2.1 Rayonnement d'un atome quantique . . . 264
3.2.2 Diffusion du rayonnement . . . 266
3.3 Diffusion par un milieu dense . . . 276
3.3.1 Notations. Champ diffusé. . . . 276
3.3.2 Cas d'un milieu homogène . . . 279
3.3.3 Diffusion par un cristal . . . 280
3.3.4 Diffusion par un milieu désordonné . . . 282
3.3.5 Influence de la dynamique du milieu . . . 283

Appendice 1 Réaction de rayonnement: modèle d'Abraham–Lorentz 287

V Electromagnétisme dans la matière 291

1 Equations de Maxwell dans la matière 295
1.1 Champs et charges microscopiques et macroscopiques . . . 295
1.2 Distributions moyennées . . . 296
1.2.1 Densités microscopiques . . . 296
1.2.2 Moments multipolaires . . . 297
1.2.3 Densités macroscopiques . . . 298
1.2.4 Expression des densités de charges liées . . . 299
1.3 Equations de Maxwell macroscopiques . . . 303
1.3.1 Déplacement électrique, Induction magnétique . . . 304
1.3.2 Considérations énergétiques . . . 305

2 Réponse linéaire **307**
2.1 Susceptibilités . . . 308
2.2 Polarisabilité électrique . . . 310
2.2.1 Polarisabilité induite . . . 310
2.2.2 Polarisablilité d'orientation . . . 310
2.3 Polarisabilités magnétiques . . . 312
2.3.1 Diamagnétisme . . . 312
2.3.2 Paramagnétisme . . . 314
2.4 Lien entre polarisabilité et susceptibilité . . . 316
2.4.1 Cas électrique . . . 316
2.4.2 Cas magnétique . . . 319
2.5 Perméabilité et permittivité relatives . . . 321
2.5.1 Définitions et équations de Maxwell . . . 321
2.5.2 Considérations énergétiques . . . 322
2.6 Relations de Kramers–Krönig . . . 324

3 Propagation dans les milieux linéaires **327**
3.1 Equations de propagation . . . 327
3.2 Milieux diélectriques . . . 328
3.3 Milieux conducteurs . . . 330
3.3.1 Modèle de conductivité . . . 330
3.3.2 Propagation très basse fréquence . . . 332
3.3.3 Propagation haute fréquence . . . 333
3.3.4 Régime intermédiaire . . . 333
3.3.5 Ondes de Plasma . . . 335
3.4 Relations de passage . . . 336
3.4.1 Conditions aux limites . . . 336
3.4.2 Passage entre deux milieux diélectriques . . . 337

Chapitre 1

Introduction générale

Le programme du module L3 du Magistère Interuniversitaire de Physique couvre essentiellement l'électromagnétisme avancé et la relativité restreinte.

Il n'est peut être pas utile de justifier longuement, à ce niveau, l'intérêt d'un cours d'électromagnétisme. Il s'agit d'une des quatre interactions fondamentales dont nous pensons qu'elles suffisent à expliquer l'ensemble des interactions observées à toutes les échelles, depuis les structures du proton ou du neutron, jusqu'à celles des amas de galaxies. Parmi ces quatre interactions, l'électromagnétisme jouit d'un statut remarquable. Il est d'abord responsable de la stabilité de l'édifice atomique, de toutes les réactions chimiques. C'est aussi souvent par l'intermédiaire d'interactions électromagnétiques que nous pouvons acquérir des informations sur le monde qui nous entoure. C'est essentiellement un problème d'électromagnétisme que d'explorer la structure du proton par des électrons de haute énergie. C'est aussi dans le domaine de l'optique, visible, infrarouge ou micro-onde, que nous pouvons explorer la structure de l'univers et remonter aux premiers stades de sa formation. Enfin, l'électromagnétisme, dans sa forme la plus achevée, l'électrodynamique quantique, est sans doute la théorie physique la mieux vérifiée et la plus sûre aujourd'hui. Les techniques modernes de théorie des champs appliquées à l'électromagnétisme permettent en effet de prédire des quantités physiquement mesurables, telles que le célèbre facteur gyromagnétique anormal de l'électron (plus connu sous le nom de $g-2$) ou les déplacements de Lamb de l'atome d'hydrogène avec des précisions pouvant atteindre 10^{-11}, limitées essentiellement à ce niveau par notre connaissance imparfaite de la structure des hadrons (protons et autres nucléons).

Il faut voir aussi, d'un point de vue plus historique, que l'électromagnétisme a joué un rôle essentiel, au début de ce siècle, dans le développement de la physique moderne. C'est en fait par ses incompatibilités avec les théories antérieures que l'électromagnétisme a contribué à renouveler totalement notre vision du monde. La première de ces incompatibilités est celle de l'électromagnétisme avec la thermodynamique classique. Quand on a essayé, à la fin du siècle dernier, de calculer à partir de la toute nouvelle théorie de Maxwell (1865) le spectre du rayonnement d'un corps noir (totalement absorbant) en équilibre thermodynamique, on s'est heurté à une difficulté en apparence insurmontable. Les lois classiques (loi de Rayleigh–Jeans par exemple), établies simplement à partir des équations de Maxwell et de considérations énergétiques, prévoient en effet un rayonnement de puissance infinie, avec un spectre divergeant aux hautes fréquences, ce qui n'est (heureusement) pas vérifié expérimentalement. Ce n'est qu'en 1900 que Planck résolut le problème en quantifiant (sans vraiment croire à une authentique nature quantique de la matière ou du rayonnement) les échanges d'énergie matière–rayonnement. En fait, la nature corpusculaire du rayonnement ne sera établie sur des arguments convaincants que par Einstein, qui analyse en 1905 les fluctuations d'un rayonnement en équilibre thermodynamique et identifie un terme similaire à celui qu'on obtient pour un gaz de particules. Il découvre ainsi le photon (le nom n'apparaîtra que bien plus tard) et interprète en ces termes les propriétés de l'effet photoélectrique. Cette idée de quantifier les grandeurs classiques devait, bien sûr, conduire ensuite à la formulation moderne de la physique quantique.

L'incompatibilité de l'électromagnétisme de Maxwell avec la cinématique classique a joué, elle aussi, un rôle essentiel qui sera largement illustré dans ce cours. Les équations de Maxwell prédisent, comme chacun sait, une propagation d'ondes électromagnétiques avec une vitesse universelle, c. La cinématique classique impliquant la loi standard de composition des vitesses, l'opinion communément répandue à la fin du siècle dernier était que cette vitesse était relative à un milieu immatériel remplissant tout l'espace, l'éther. Ce milieu n'a pas tardé à poser quelques problèmes. Il fallait d'abord qu'il soit pratiquement immatériel, pour se laisser traverser sans friction apparente par les planètes. Il fallait en même temps qu'il soit extrêmement rigide pour transmettre des vibrations transverses à grande vitesse. Plus encore, cet éther posait des problèmes d'ordre plus philosophique, en réintroduisant un référentiel absolu. Enfin, l'hypothèse de l'éther s'effondra tout à fait quand les expériences de Michelson, justement célèbres, montrèrent que l'éther semblait immobile par rapport à la terre. A moins d'en revenir à un anthropocentrisme intolérable ou d'inventer des modifications ad hoc complètement artificielles de la théorie (entraînement de l'éther par les masses en mouvement, par exemple), il n'y avait plus comme issue que d'inventer la relativité restreinte (en 1905) en renouvelant complètement les bases de la cinématique et de la dynamique, avec des conséquences philosophiques importantes (abandon de l'universalité du temps), puis la relativité générale, qui donne de la gravitation une interprétation complètement géométrique. Il est assez remarquable, d'ailleurs, que les deux incompatibilités que nous venons de discuter aient conduit à deux théories (relativité générale et mécanique quantique) parfaitement vérifiées dans la limite des expériences actuelles mais encore incompatibles, en dépit des efforts de générations de physiciens.

La dernier rôle historiquement important de l'électromagnétisme est plus récent. Dès la formulation de la mécanique quantique moderne, au début des années 30, on a tenté de quantifier le champ électromagnétique et de retrouver ainsi rigoureusement la quantification heuristique de Planck. Si tout se passe bien avec les procédés de quantification standard (le terme adéquat est "canonique") quand on ne considère qu'un mode du rayonnement (une seule onde plane, par exemple), les choses se gâtent quand on veut tenir compte de toutes les fréquences et de toutes les directions de propagation. Il apparaît alors des infinis dans le calcul de toute quantité physique. Ce n'est que relativement récemment (1947 environ) qu'on a pu se débarrasser systématiquement de ces infinis. Ces techniques, en particulier la renormalisation, mises au point pour l'électromagnétisme, ont depuis joué un rôle essentiel dans la théorie des champs, puisqu'on ne sait pratiquement, encore aujourd'hui, traiter que les théories renormalisables.

Il n'est bien entendu pas possible d'aborder ces problèmes dans un cours de licence. Nous nous contenterons d'étudier certaines propriétés de l'électromagnétisme classique (de Maxwell) qui ne sont que très partiellement abordées dans les classes "élémentaires" et de discuter des liens profonds entre électromagnétisme et relativité restreinte. A ce programme relativement ambitieux, il a été décidé récemment d'ajouter un bref chapitre de mécanique analytique. Il s'agit en effet d'une formulation élégante et puissante de la dynamique newtonienne classique, qui nous sera fort utile, dans le cours de relativité restreinte, pour montrer que l'électromagnétisme de Maxwell est en fait une des dynamiques relativistes les plus simples que l'on puisse construire avec une interaction champ–matière non triviale. Cette introduction sera également utile pour éclairer le cours de physique statistique classique, dans laquelle la fonction de Hamilton de la mécanique classique joue un rôle essentiel. Enfin les formulations lagrangiennes et hamiltoniennes de la mécanique classique jouent un rôle essentiel en mécanique quantique, en fournissant les techniques nécessaires pour un passage rigoureux d'une théorie classique à la théorie quantique correspondante (la fameuse quantification canonique évoquée plus haut). Certains des objets du formalisme de la mécanique quantique correspondent d'ailleurs de très près à des analogues en mécanique analytique. Les commutateurs ne sont que la version matricielle des crochets de Poisson que nous introduirons dans cette première partie sur la mécanique analytique. Nous conclurons cette partie par un appendice sur le modèle de Bohr de la structure atomique. Il s'agit en effet d'un modèle, fondé sur des concepts de mécanique analytique, qui nous sera fort utile pour des discussions qualitatives dans la suite du cours.

La deuxième partie du cours sera consacrée à la relativité restreinte. Nous essaierons de montrer pourquoi l'électromagnétisme de Maxwell est incompatible avec la cinématique de Galilée ou de Newton. Nous construirons alors, en nous fondant sur des hypothèses très simples et naturelles, une nouvelle cinématique. Nous devrons donc reconstruire aussi la dynamique des particules matérielles. Nous chercherons alors à construire une théorie décrivant l'interaction de particules matérielles par l'intermédiaire d'un champ. Nous prendrons la forme la plus simple possible pour les fonctions de Lagrange décrivant ce champ et son interaction avec la matière. En utilisant les résultats de mécanique analytique, nous montrerons alors que la théorie ainsi construite n'est autre que l'électromagnétisme de Maxwell! Nous aurons ainsi bouclé la boucle mais montré surtout que la formulation de Maxwell, arrivée 40 ans avant la relativité, est naturellement relativiste. Nous obtiendrons enfin, en utilisant cette approche relativiste, un certain nombre de résultats de pur électromagnétisme, en particulier sur les bilans d'énergie–impulsion du champ, particulièrement pénibles à obtenir par d'autre méthodes.

La troisième partie du cours sera consacrée aux phénomènes de propagation et de diffraction des champs électromagnétiques. Nous donnerons en particulier explicitement la solution des équations de Maxwell en termes de potentiels retardés. Cette démonstration, outre son importance, fait intervenir la technique très puissante des fonctions de Green, qui sont d'un usage courant dans de nombreux domaines de la physique et qui jouent un rôle essentiel dans l'établissement de la théorie rigoureuse de la diffraction. Le deuxième chapitre de cette partie sera donc consacré à une discussion détaillée de la théorie rigoureuse de la diffraction. Nous montrerons quelle est la démarche pour passer des équations de Maxwell à l'approximation de Fraunhofer de la diffraction paraxiale, telle qu'elle est enseignée dans les classes élémentaires. Le troisième chapitre de cette partie sera consacrée à une discussion rapide et très qualitative de quelques applications de la diffraction dans le domaine de traitement optique du signal. Ce chapitre sera suivi de quatre appendices qui pourront être ignorés en première lecture. Le premier traitera les problèmes de choix de jauge. Si la relativité impose une jauge, il en est d'autres qui peuvent être utiles pour des problèmes où l'invariance relativiste peut être temporairement masquée. Dans le second, nous explorerons l'analogie formelle entre le rayonnement et l'oscillateur harmonique, en introduisant les variables normales du champ. C'est là un étape essentielle vers la quantification du rayonnement, que nous évoquerons très brièvement et qualitativement. Le troisième, en application directe des principes de la diffraction, traitera des faisceaux gaussiens, essentiels en optique laser. Enfin, le quatrième montrera comment l'optique géométrique peut être déduite de l'électromagnétisme quand on ne s'intéresse qu'à des champs variant lentement à l'échelle de la longueur d'onde.

Nous nous pencherons ensuite, dans la quatrième partie, sur le problème du calcul des champs rayonnés par divers types de sources. Si la solution en termes de potentiels retardés est parfaitement explicite, elle n'est guère manipulable dans la plupart des cas. Nous nous occuperons essentiellement dans ce chapitre de trois types de sources. Nous commencerons par examiner le rayonnement produit par une charge en mouvement (éventuellement relativiste) imposé. Nous pourrons ainsi nous pencher sur le problème du rayonnement de freinage et de la réaction de rayonnement essentiels dans la description des accélérateurs de particules et dans celle de l'interaction de particules chargées énergétiques avec la matière. Nous pourrons aussi traiter le rayonnement du dipôle électromagnétique, constitué d'une simple charge oscillant de façon sinusoïdale au voisinage de l'origine. En raison de l'importance de ce cas, nous expliciterons le calcul du champ à des distances arbitraires. Nous examinerons, dans le deuxième chapitre, des répartitions de courants classiques oscillants (des antennes) que nous traiterons par la technique des développements multipolaires, en étudiant en détail les dipôles et quadripôles électriques ainsi que les dipôles magnétiques, qui joueront un rôle essentiel dans la partie suivante. Nous nous pencherons enfin, dans le dernier chapitre, sur le rayonnement de sources atomiques et nous étudierons en particulier la diffusion de la lumière par un atome unique. Nous montrerons, en les comparant explicitement, qu'un modèle classique très simple donne des résultats qualitativement très comparables à celui d'un modèle quantique beaucoup plus rigoureux. Nous montrerons néanmoins où sont les limites de ce traitement ignorant le caractère quantique du champ, en discutant en particulier d'expériences récentes.

La cinquième partie sera consacrée aux champs électromagnétiques dans la matière. Nous montrerons d'abord comment on peut se débarrasser, par des procédures de moyennage appropriées et l'introduction de nouveaux champs, de la formidable complexité des répartitions de charges dans la matière dense. Nous verrons comment la théorie de la réponse linéaire permet d'éliminer de ces champs supplémentaires. Nous l'aborderons ici pour la première fois les méthodes de réponse linéaires applicables, sous des formes diverses, à des domaines très variés, de la mécanique quantique au traitement du signal. Nous nous pencherons alors sur la notion de susceptibilité, qui décrit tout autant les phénomènes dispersifs dans le milieux transparents (l'indice de réfraction), que les échanges d'énergie. Nous montrerons comment la causalité introduit des relations très belles et très fondamentales entre dispersion et absorption. Là encore, ces relations sont d'un champ d'application beaucoup plus large que l'électromagnétisme qui nous fournira une première occasion de les aborder.

Nous supposerons connues dans ce polycopié et dans le cours, un certain nombre de notions.

- **Mécanique du point:** notion de vitesse, accélération, référentiel, changement de référentiel galiléen, principe fondamental, énergies cinétiques et potentielles, moment cinétique.
- **Electrostatique:** champ, potentiel, théorème de Gauss, utilisation des propriétés de symétrie, énergie électrostatique. Notions d'électrostatique des conducteurs.
- **Magnétostatique:** champ, potentiel vecteur, théorème d'Ampère, utilisation des propriétés de symétrie, énergie magnétostatique.
- **Electrodynamique:** équations de Maxwell, conditions de Jauge, propagation, notion d'onde plane, polarisation, potentiels retardés, énergétique des champs électromagnétiques dans le vide (densité d'énergie et vecteur de Poynting). Quelques notions sur l'électrodynamique des milieux matériels
- **Optique:** quelques notions élémentaires d'optique géométrique, interférences et diffraction dans la limite de Fraunhofer.
- **Mathématiques:** calcul vectoriel, analyse vectorielle (gradient, divergence, rotationnel...). intégration, différents systèmes de coordonnées (cartésien, cylindrique, sphérique), bases d'algèbre linéaire, équations différentielles élémentaires. Séries de Fourier et transformées de Fourier

Ce polycopié est en fait relativement plus complet que le cours lui même, dont la durée limitée ne permet pas de traiter en détails tous les sujets énumérés ici. Pour approfondir encore d'avantage le sujet, on pourra recourir à de nombreux manuels. Pour ce qui est de la mécanique analytique, nous recommandons le Landau (Mécanique), très sec mais très complet, et le Goldstein (Mécanique classique) que l'on peut trouver en versions anglaise et traduite. C'est un livre très (trop?) complet. Il est de loin préférable de lire une édition récente, les anciennes étant un peu poussiéreuses. Pour la relativité, il existe une infinité de manuels. On pourra se référer, là encore au Landau (théorie des champs) si on n'est pas rebuté par le style de cet ouvrage et les notations, un peu anciennes. Il n'est pas inutile non plus de regarder les articles originaux d'Einstein. Un article de revue de 1907, en particulier, que l'on trouvera traduit dans la récente édition d'une sélection d'articles (édition Einstein, Relativités I, Seuil CNRS), est un modèle de pédagogie et ferait un excellent manuel.

Pour tout ce qui concerne les problèmes d'électromagnétisme et aussi pour la relativité il est indispensable d'avoir au moins parcouru le Jackson (Classical Electrodynamics). Ce très beau et très gros livre est la bible du domaine. Il est extrêmement exhaustif et d'une lecture suffisamment facile (surtout les éditions récentes). En fait, il pourrait à lui seul remplacer 80% de ce polycopié, dont certains chapitre sont fortement inspirés. Son seul défaut est l'utilisation exclusive du système d'unités CGS/UES, ce qui fait que les équations ne sont que difficilement reconnaissables pour des européens habitués au système dit international. Fort heureusement, Jackson fournit, dans un appendice très intelligemment conçu, les règles de transformation qui sont finalement assez simples. Pour l'optique

(diffraction, aspects ondulatoires, problèmes de cohérence, de polarisation) on pourra consulter avec profit le Born et Wolf, vieux manuel ennuyeux mais extraordinairement exhaustif.

Remerciements

Ce polycopié doit beaucoup, pour la partie de pur électromagnétisme, à un polycopié rédigé il y a quelques années par S. Haroche, pour ce qui était alors le cours M3 du MIP. Il doit aussi énormément aux "caïmans" qui m'ont assisté dans les dernières années: M. Benamar, M.C. Angonin, J.M. Daul, C. Dupraz, L. Rezeau et J. Hare. Ce dernier, en particulier, a consacré beaucoup de temps à la relecture attentive du manuscrit et a suggéré de nombreuses améliorations. Je remercie également E. Reyssat, élève de la promotion 1999, qui a relevé de nombreuses erreurs typographiques.

Partie I

Mécanique analytique

Introduction

La mécanique analytique n'apporte rien de conceptuellement nouveau par rapport aux formulations standard de la dynamique newtonienne (principe fondamental, théorème de l'énergie cinétique et autres points marquants de l'enseignement élémentaire de la mécanique), mais en constitue une formulation très élégante. Parfaitement adaptée à la description de systèmes où les mouvements sont sujets à des contraintes (un cauchemar avec les formulations "standard"), à l'utilisation de techniques de perturbations, ce qui explique son succès toujours certain auprès des astronomes, elle est souvent d'un usage infiniment plus pratique que les formulations plus élémentaires.

Il s'agit aussi d'un cas particulier d'une approche très fructueuse dans des domaines variés de la physique: une méthode variationnelle. En mécanique analytique, nous ne préciserons pas les équations locales que doit vérifier à chaque instant le mouvement de la particule. Nous donnerons en fait une condition prescrivant à une intégrale portant sur l'ensemble du mouvement d'être extrémale. Parmi toute les trajectoires permises par la cinématique, mais parfois absurdes pour la dynamique, il nous faudra choisir la bonne en respectant cette règle. En fait, la description du mouvement en mécanique analytique est très semblable à la description des rayons lumineux avec le principe de Fermat. Là aussi, on doit choisir parmi tous les trajets possibles celui qui rend extrémale une intégrale qui n'est autre que la durée du trajet.

Surtout, et bien qu'il s'agisse d'un formalisme datant, avec Lagrange et Hamilton, de la fin du XVIIIème ou du XIXème siècle, elle est parfaitement adaptée aux approches modernes de la physique. Elle joue ainsi un rôle essentiel en mécanique statistique, elle est à l'origine de la quantification des dynamiques classiques, elle est fortement apparentée aux formulations modernes de la mécanique quantique en termes d'intégrales de chemin. Elle nous sera enfin d'une grande utilité pour reconstruire l'électromagnétisme à partir de la relativité.

Cette partie se compose de deux chapitres principaux. Dans le premier, qui sera le plus étoffé, nous donnons la formulation lagrangienne de la mécanique analytique, qui est celle que nous utiliserons dans la partie de relativité. Nous insisterons sur la notion de coordonnée généralisée, qui permet de traiter de façon naturelle les contraintes et nous examinerons comment on peut incorporer dans le formalisme un certain nombre d'interactions. Un point important dans ce domaine sera l'établissement de la fonction de Lagrange pour des particules chargées en interaction avec un champ, dont nous montrerons qu'elle redonne bien le force de Lorentz. Enfin, nous déduirons d'un certain nombre de symétries fondamentales de la nature (invariance dans le temps, dans l'espace, invariance par rotation) les lois de conservation essentielles (énergie, impulsion, moment cinétique). Cette approche qui lie les lois de conservation aux propriétés de symétrie est en fait très générale et très puissante.

Le deuxième chapitre sera consacré à une brève revue du formalisme hamiltonien. Brève parce que le sujet est extrêmement vaste, en particulier en ce qui concerne les transformations canoniques et les liens avec la mécanique quantique, brève aussi parce que le formalisme hamiltonien ne sera guère utilisé en grand détail dans les cours de première année.

Chapitre 1

Formulation lagrangienne

1.1 Description du système: coordonnées généralisées

Nous considérerons donc un système composé de N particules matérielles repérées par un indice grec α, variant de 1 à N. Une telle description peut convenir à tout système discret de particules ponctuelles mais aussi à la description du mouvement d'un solide, après une discrétisation convenable en éléments infinitésimaux. Les masses, charges électriques, positions, vitesses et accélérations des particules seront dénotées respectivement m_α, q_α, $\mathbf{r}_\alpha$, $\mathbf{v}_\alpha = \dot{\mathbf{r}}_\alpha$, $\mathbf{a}_\alpha = \dot{\mathbf{v}}_\alpha = \ddot{\mathbf{r}}_\alpha$ (nous désignerons souvent dans la suite les dérivées temporelles par des symboles pointés. Les caractères gras représentent des quantités vectorielles).

L'approche standard de la mécanique newtonienne est alors d'écrire le principe fondamental de la dynamique, reliant les accélérations des diverses particules constituant le système aux forces s'exerçant sur elles. L'expression de ces forces est donnée, en fonction de la configuration du système, soit par des lois fondamentales (force de Lorentz, par exemple), soit par des lois phénoménologiques (forces de frottement...). Par exemple, dans le cas de particules en interaction électromagnétique, on écrirait:

$$m_\alpha \mathbf{a}_\alpha = \mathbf{f}_\alpha = q_\alpha(\mathbf{E}(\mathbf{r}_\alpha) + \mathbf{v}_\alpha \times \mathbf{B}(\mathbf{r}_\alpha)) \,, \tag{1.1}$$

où $\mathbf{E}$ et $\mathbf{B}$ sont les champs électrique et magnétique déterminés, en fonction des positions de particules, par la solution des équations de Maxwell.

Si l'écriture de toutes les équations dynamiques du système permet en principe, en y ajoutant les conditions initiales convenables, de déterminer complètement le mouvement, cette résolution peut être très délicate. C'est en particulier le cas quand il existe des contraintes: les positions (ou les vitesses) des particules doivent constamment obéir à un certain nombre de relations. Imaginons, par exemple, le cas de deux pendules accrochés l'un à l'extrémité de l'autre et contraints à se déplacer dans un plan (voir figure 1.1). Dans les formulations classiques, on doit associer à ces différentes liaisons des forces (force de tension des fils constituant les pendules, force de réaction du support commun...). Ces forces sont de nouvelles inconnues dans le problème qui doivent être déterminées en même temps que les variables dynamiques intéressantes. Bien entendu, elles compliquent beaucoup la résolution du problème.

L'idée de la mécanique analytique est de se débarrasser de ces forces inconnues en n'employant que des coordonnées indépendantes qui ne seront soumises à aucune contrainte. Nous les appellerons "coordonnées généralisées". Ces coordonnées sont de nature arbitraire (des positions, des angles...) mais doivent déterminer de façon univoque l'état mécanique du système si on prend en compte les contraintes. On pourra déterminer le mouvement en écrivant une équation différentielle pour chacune de ces coordonnées. Considérons, pour fixer les idées, le cas du double pendule. Il y a a priori six paramètres pour décrire le système (les positions des deux masses). En fait, les contraintes diminuent considérablement la dimensionnalité du problème. D'abord, les fils sont de longueur constante, soit

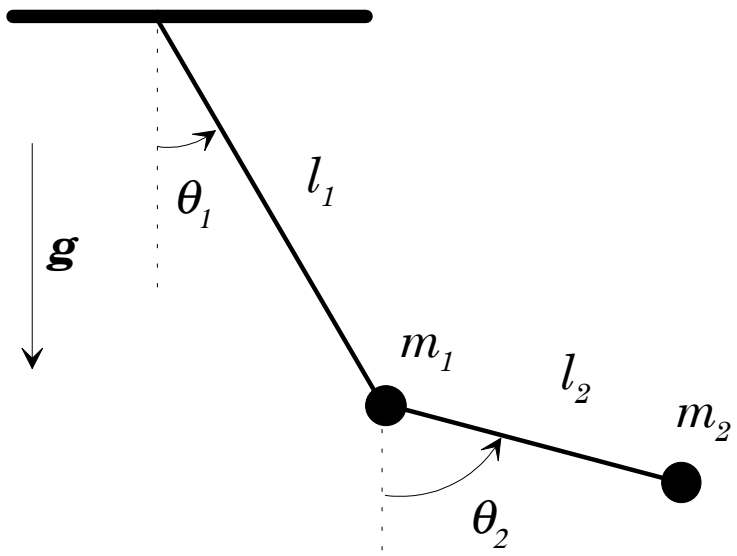

Figure 1.1: Un problème classique de mécanique: deux pendules liés, asservis à se déplacer dans un plan. Les angles θ_1 et θ_2 suffisent à décrire complètement l'état mécanique du système

deux relations. Ensuite, le mouvement s'effectue dans un plan, ce qui fournit encore deux relations (par exemple en écrivant que le produit scalaire de la position avec la normale au plan est nul). Il n'y a donc en fait que deux variables indépendantes qui décrivent le mouvement. Un choix tout à fait naturel de coordonnées généralisées dans ce cas est de prendre les deux angles θ_1 et θ_2 des pendules avec la verticale.

Plus généralement, nous supposerons que les liaisons entre les simples coordonnées cartésiennes des particules sont holonomes: il existe $3N - n$ relations du type $f_j(\mathbf{r}_\alpha) = 0$. De telles relations décrivent convenablement toutes les contraintes directes entre coordonnées, à condition qu'elles soient indépendantes du temps (comme celles que nous venons de voir, si la longueur des fils des pendules est invariable)[1]. Elles ne décrivent pas les contraintes entre vitesses (par exemple le roulement sans glissement) mais nous verrons plus loin comment on peut en tenir compte. Il ne reste alors que n coordonnées généralisées indépendantes que nous noterons q_i, $i = 1...n$. Soulignons une fois de plus que ces coordonnées ne sont pas nécessairement cartésiennes et n'ont même pas forcément la dimension d'une longueur. Avec des relations holonomes, les positions $\mathbf{r}_\alpha$ ne dépendent que des n coordonnées généralisées et ne dépendent ni des vitesses ni du temps explicitement. Il nous faut maintenant donner les lois permettant d'établir les n équations différentielles déterminant la dynamique des q_i.

[1]En fait, la plupart des résultats que nous établirons dans ce chapitre seraient également valables si les relations faisaient intervenir une dépendance explicite en temps, sous la forme $f_j(\mathbf{r}_\alpha, t) = 0$. Les positions dépendraient alors des coordonnées généralisées, mais présenteraient aussi une dépendance explicite en temps. L'énergie cinétique, par exemple, qui est une forme quadratique des dérivées des coordonnées généralisées dans le cas habituel, ferait intervenir des termes linéaires dans ces dérivées, ou même des termes n'en dépendant pas. Nous préciserons, si nécessaire, quels sont les résultats qui dépendent de façon critique de cette hypothèse.

1.2 Principe de moindre action

1.2.1 Enoncé

On postule qu'il existe une fonction $L(q_i, \dot{q}_i, t)$, dite fonction de Lagrange ou lagrangien, homogène à une énergie[2], qui est telle que l'action

$$S = \int_{t_1}^{t_2} L(q_i, \dot{q}_i, t)\, dt \ , \tag{1.2}$$

soit *extrémale* pour la trajectoire effectivement suivie par le système de t_1 à t_2 entre $q_i(1)$ et $q_i(2)$, valeurs initiales et finales des coordonnées généralisées.

Ce que nous postulons ainsi n'est pas directement un ensemble d'équations différentielles pour les variables dynamiques (que nous pourrons déduire et dont nous montrerons qu'elles sont équivalentes aux formulations standard de la mécanique). Nous nous donnons plutôt un principe variationnel qui postule le caractère extrémal d'une certaine intégrale calculée sur la trajectoire en fonction de celle-ci. Il y a de nombreux autres exemples de principes variationnels en physique. Les lois de l'optique géométrique, par exemple, peuvent se déduire du principe de Fermat qui postule que le rayon lumineux effectivement suivi réalise un extremum (en général un minimum) du temps de parcours.

Le fait que nous prenions une fonction de Lagrange ne dépendant que des positions et des vitesses (mais pas de dérivées d'ordre supérieur) exprime, comme nous le verrons, que les équations fondamentales de la dynamique sont d'ordre deux par rapport au temps. D'autre part, nous spécifions les deux conditions "initiales" nécessaires pour chaque coordonnée en donnant les positions initiales et finales et non les positions et vitesses initiales. Si ces deux formulations sont bien sûr équivalentes, la première est plus avantageuse pour varier l'action sur toutes les trajectoires possibles entre deux points.

Bien sûr, un principe variationnel est d'emploi moins commode en pratique qu'un ensemble d'équations différentielles. Il faut, en principe, imaginer toutes les trajectoires possibles (continues et dérivables) entre les conditions initiales et finales, déterminer l'action sur chacune et déterminer celles qui rendent l'action extrémale. Nous verrons, dans le prochain paragraphe, comment en déduire un système d'équations différentielles beaucoup plus commodes.

1.2.2 Equations de Lagrange

Nous considérons donc deux trajectoires possibles entre $q(1)$ et $q(2)$. L'une, que nous noterons simplement $q_i(t)$, est la trajectoire effectivement suivie. L'autre que nous appellerons "trajectoire variée", infiniment proche, correspond à chaque instant aux positions $q_i(t) + \delta q_i(t)$, où $\delta q_i(t)$ est un accroissement infinitésimal de la position (voir figure 1.2). Ces deux trajectoires doivent obéir aux mêmes conditions initiales et finales. On a donc $\delta q(1) = \delta q(2) = 0$. Nous supposerons que les q_i et δq_i sont deux fois différentiables. Le fait que les q_i donnent la trajectoire effectivement suivie a pour conséquence que l'action S est extrémale sur cette trajectoire et ne varie donc pas au premier ordre dans les δq_i quand on passe à la trajectoire variée. Or la variation de l'action s'écrit simplement:

$$\delta S = \int_{t_1}^{t_2} \left(L(q_i + \delta q_i, \dot{q}_i + \dot{\delta q}_i, t) - L(q_i, \dot{q}_i, t)\right) dt. \tag{1.3}$$

En développant L au premier ordre dans les δq_i, on a:

$$\delta S = \sum_i \int_{t_1}^{t_2} \frac{\partial L}{\partial q_i} \delta q_i \, dt \ + \sum_i \int_{t_1}^{t_2} \frac{\partial L}{\partial \dot{q}_i} \frac{d\delta q_i}{dt} \, dt \ . \tag{1.4}$$

[2]et que l'on saura écrire si on connaît la nature des forces qui s'exercent sur le système

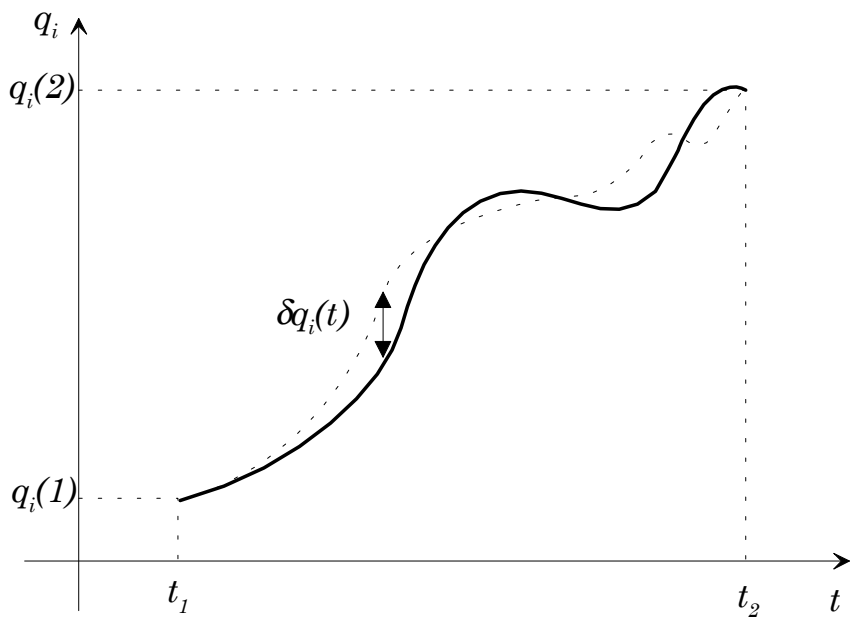

Figure 1.2: Trajectoire effectivement suivie (ligne continue) et trajectoire variée (pointillée). La trajectoire variée s'écarte infinitésimalement de la trajectoire effectivement suivie et coïncide avec celle-ci aux extrémités.

Pour fixer la trajectoire, nous recherchons une condition sur les δq_i. Pour éliminer leurs dérivées temporelles dans l'expression précédente nous intégrons les termes de la seconde somme par parties. On obtient alors:

$$\delta S = \sum_i \int_{t_1}^{t_2} \left[\frac{\partial L}{\partial q_i} - \frac{d}{dt}\frac{\partial L}{\partial \dot{q}_i}\right] \delta q_i \, dt \; + \left[\frac{\partial L}{\partial \dot{q}_i}\delta q_i\right]_{t_1}^{t_2} . \tag{1.5}$$

Les accroissements infinitésimaux δq_i s'annulant aux extrémités de la trajectoire, le terme tout intégré est identiquement nul. La somme, elle, ne peut s'annuler pour des δq_i arbitraires (et indépendants) que si les n équations différentielles:

$$\frac{\partial L}{\partial q_i} - \frac{d}{dt}\frac{\partial L}{\partial \dot{q}_i} = 0 \qquad (1 \leq i \leq n) \tag{1.6}$$

sont simultanément vérifiées. Ce système différentiel avec les conditions aux limites (fournissant $2n$ conditions indépendantes) détermine complètement les n coordonnées généralisées indépendantes. Ces équations différentielles sont du second ordre par rapport au temps puisque L dépend a priori des $\dot{q}_i$. Quelques remarques s'imposent à ce point.

- Les équations du mouvement ne changent pas quand L est multipliée par une constante. Cette liberté correspond seulement à un choix d'unités (L a, rappelons–le, la dimension d'une énergie).

- La structure de L doit obéir aux symétries du système physique (invariance par translation dans le temps, dans l'espace..). Nous verrons, dans les prochains paragraphes, que cela conduit à des conséquences importantes en termes de lois de conservation.

- Les équations du mouvement sont inchangées si on ajoute à L la dérivée totale par rapport au temps d'une fonction des coordonnées et du temps. Posons en effet $L' = L + df(q_i, t)/dt$ (la fonction f que nous dérivons ne doit dépendre que des q_i pour que le lagrangien modifié ne dépende que des q_i et $\dot{q}_i$). L'action S', calculée avec la nouvelle fonction de Lagrange, ne diffère de S que par un terme de la forme $[f]_{t_1}^{t_2}$, qui ne dépend manifestement pas de la trajectoire suivie entre 1 et 2. On pourra vérifier par simple substitution que les équations de Lagrange ne

changent pas quand on effectue cette modification du lagrangien. Cette liberté (qui n'est pas sans évoquer la liberté de jauge en électromagnétisme) est parfois fort utile pour simplifier la forme du lagrangien.

- La fonction de Lagrange est additive. Considérons en effet deux systèmes physiques indépendants et n'interagissant pas, décrits par les coordonnées généralisées q_1 et q_2 et par les fonctions de Lagrange L_1 et L_2. Le système global est défini par la réunion des deux ensembles de coordonnées généralisées et il est évident que la fonction de Lagrange associée au système complet est $L = L_1 + L_2$.

- Notons enfin que, dans le cas très simple où les coordonnées généralisées coïncident avec les coordonnées cartésiennes des particules individuelles (ce qui est par exemple le cas quand il n'y a aucune contrainte sur le mouvement), les équations de Lagrange peuvent s'écrire:

$$\boldsymbol{\nabla}_{\mathbf{r}_\alpha} L = \frac{d\boldsymbol{\nabla}_{\mathbf{v}_\alpha} L}{dt} \ , \tag{1.7}$$

en faisant intervenir les opérateurs gradient par rapport aux positions et vitesses de chaque particule.

Il nous reste maintenant, pour que ce formalisme ait un sens, à donner la forme de la fonction de Lagrange en fonction des interactions que subissent les particules.

1.3 Expressions de la fonction de Lagrange

Cette section est essentielle dans ce chapitre, puisqu'elle nous permettra de traiter effectivement des problèmes de mécanique par le formalisme lagrangien. Nous nous pencherons d'abord sur le cas de la particule libre, puis sur le cas de particules interagissant par des forces conservatives (dérivant d'un potentiel), sur le cas de particules soumises à des forces extérieures au système (avec quelques applications au cas important du mouvement dans le champ de pesanteur) et enfin sur le cas de particules en interaction avec un champ électromagnétique.

1.3.1 Particule unique libre

Ce cas élémentaire n'a de mérite que pédagogique. Il est en effet évident dès l'abord que toute fonction de Lagrange conduisant à des équations s'écrivant $\dot{\mathbf{v}} = 0$ conviendra. Une simple fonction proportionnelle à v^2 vérifie (entre autres) cette propriété. La fonction de Lagrange d'une particule libre unique est donc, à un choix d'unités près, identique à l'énergie cinétique. Nous allons toutefois montrer comment on peut arriver à ce résultat en utilisant les propriétés d'invariance et de symétrie.

Les coordonnées généralisées coïncident dans ce cas avec les coordonnées cartésiennes standard. La particule étant libre, L ne peut explicitement dépendre du temps (invariance par translation dans le temps), ni de la position $\mathbf{r}$ de la particule (invariance par translation spatiale), ni enfin de la direction de sa vitesse $\mathbf{v}$ (invariance par rotation). L doit donc être une fonction du carré du module de la vitesse: $L = f(v^2)$.

Les équations de Lagrange se résument alors à:

$$\boldsymbol{\nabla}_{\mathbf{r}} L = 0 = \frac{d}{dt}\left[\frac{df}{dv^2}\boldsymbol{\nabla}_{\mathbf{v}} v^2\right] \ , \tag{1.8}$$

qui conduisent bien évidemment, à moins que f ne soit une constante, à $\dot{\mathbf{v}} = 0$ et donc à un mouvement rectiligne uniforme.

Pour préciser davantage la forme de L et de f, il nous faut ajouter une condition supplémentaire: le résultat précédent doit être invariant dans un changement de référentiel galiléen. Considérons pour

cela un référentiel $\mathcal{R}$, dans lequel la fonction de Lagrange est $L = f(v^2)$ et un référentiel $\mathcal{R}'$ en translation uniforme par rapport à $\mathcal{R}$ avec une vitesse infinitésimale $\boldsymbol{\varepsilon}$. La fonction de Lagrange L' dans $\mathcal{R}'$ doit s'écrire $L' = f(v'^2) = f((\mathbf{v} + \boldsymbol{\varepsilon})^2)$ (la fonction f devant manifestement être la même pour tous les référentiels). En développant au premier ordre en $\boldsymbol{\varepsilon}$, on a: $L' = L + (df/dv^2)2\mathbf{v} \cdot \boldsymbol{\varepsilon}$. Les équations du mouvement dans $\mathcal{R}$ et $\mathcal{R}'$ coïncideront si L et L' ne diffèrent que par une dérivée totale par rapport au temps. Le cas le plus simple où cela se vérifie est quand df/dv^2 est indépendant de v (la dérivée par rapport au temps étant alors simplement $d(2(df/dv^2)\mathbf{r} \cdot \boldsymbol{\varepsilon})/dt$. Le choix le plus simple est donc que f soit simplement proportionnelle à v^2. Nous poserons donc:

$$L = \frac{1}{2}mv^2 \tag{1.9}$$

et appellerons évidemment "masse de la particule" le coefficient m. Ce coefficient doit être positif. En effet, l'extremum de l'action correspondant à la propagation en ligne droite à vitesse constante est alors un minimum.

Bien sûr, ce raisonnement n'est pas d'une grande rigueur et repose largement sur le critère de simplicité pour identifier complètement la forme du lagrangien. Il illustre en revanche le genre de démarche qu'on doit effectuer pour déterminer la forme du lagrangien correspondant à un nouvelle interaction: respecter d'abord les grandes propriétés de symétrie, respecter les règles de la relativité galiléenne et enfin chercher la forme la plus simple en cas d'ambiguïté. C'est, avec quelques adaptations, la démarche que nous utiliserons plus tard pour déterminer, en relativité, les lagrangiens correspondant à l'interaction électromagnétique.

1.3.2 Système de particules interagissant par des forces dérivant d'un potentiel

Nous supposerons d'abord, pour fixer les idées, que les coordonnées généralisées coïncident avec les coordonnées cartésiennes. Nous supposerons que la force s'exerçant sur la particule α peut s'écrire $\mathbf{F}_\alpha = -\boldsymbol{\nabla}_{\mathbf{r}_\alpha} U(\mathbf{r}_1, ..., \mathbf{r}_n, t)$, où la fonction U est une énergie potentielle dépendant a priori de la position de toutes les particules dans le système. Pour ne pas restreindre la généralité, nous permettrons au potentiel de dépendre explicitement du temps. Nous pourrons ainsi traiter, par exemple, le mouvement dans des champs extérieurs variables. Nous chercherons simplement la forme de la fonction de Lagrange qui redonne les équations dynamiques habituelles.

Si U est identiquement nulle, la fonction L se résume à l'énergie cinétique totale: $L = T$ avec $T = \sum_\alpha (1/2) m_\alpha v_\alpha^2$ (ce qui se déduit évidemment du paragraphe précédent et de l'additivité de la fonction de Lagrange pour des systèmes sans interaction mutuelle). Nous vérifierons maintenant que le choix $L = T - U$ donne, pour les équations de Lagrange, les équations standard. En effet, les équations de Lagrange s'écrivent:

$$\boldsymbol{\nabla}_{\mathbf{r}_\alpha} L = \frac{d\boldsymbol{\nabla}_{\mathbf{v}_\alpha} L}{dt} \ , \tag{1.10}$$

et on a

$$\boldsymbol{\nabla}_{\mathbf{r}_\alpha} L = -\boldsymbol{\nabla}_{\mathbf{r}_\alpha} U = \mathbf{F}_\alpha \ , \tag{1.11}$$

et

$$\frac{d\boldsymbol{\nabla}_{\mathbf{v}_\alpha} L}{dt} = \frac{dm_\alpha \mathbf{v}_\alpha}{dt} = m_\alpha \mathbf{a}_\alpha = \mathbf{F}_\alpha \ . \tag{1.12}$$

Cette forme du lagrangien redonne donc bien le principe fondamental de la dynamique tel que nous le connaissions.

Considérons maintenant le cas où le système doit être décrit par des coordonnées généralisées qui ne coïncident pas avec les coordonnées cartésiennes. La fonction de Lagrange elle-même ne doit pas dépendre du choix particulier du système de coordonnées généralisées. Elle doit toujours coïncider avec la différence $T-U$ des énergies cinétiques et potentielles. Pour pouvoir écrire les équations de Lagrange, il faut exprimer ces quantités en fonction des coordonnées généralisées q_i et de leurs dérivées. C'est

toujours possible, puisque les q_i doivent déterminer de façon univoque l'état mécanique du système – quand on prend en compte les contraintes. En inversant ces relations et en les reportant dans les expressions de T et U en fonction des coordonnées cartésiennes on obtient facilement le résultat cherché.

Nous avons supposé que les positions $\mathbf{r}_\alpha$ ne dépendent que des q_i, mais pas de leurs dérivées ni du temps (ce ne serait pas le cas si les liaisons faisaient intervenir une dépendance explicite en temps). On peut donc écire: $\mathbf{v}_\alpha = \sum_i \dot{q}_i \partial \mathbf{r}_\alpha / \partial q_i$. En substituant dans $T = \sum_\alpha m_\alpha \mathbf{v}_\alpha^2/2$, on trouve en général T comme une forme quadratique définie positive des $\dot{q}_i$, dont les coefficients peuvent dépendre des q_i: $T = \sum_{i,j} A_{i,j}(q_k)\dot{q}_i\dot{q}_j$. Par exemple, on a une expression de ce genre quand on utilise les coordonnées cylindriques pour décrire le mouvement d'une particule unique $T = m(\dot{r}^2 + r^2\dot{\theta}^2 + \dot{z}^2)/2$. Notons enfin que, dans les mêmes conditions, U s'exprime simplement comme une fonction des q_i.

1.3.3 Système de particules soumises à des forces extérieures

C'est un cas particulièrement important en mécanique, puisqu'il permet, entre autres, de traiter de mouvements dans le champ de gravitation terrestre. Nous considérons donc un système S dont la dynamique est décrite par n coordonnées généralisées q_i et dont nous cherchons le mouvement. Les particules de S interagissent entre elles par des forces dérivant d'une énergie potentielle et interagissent avec les particules d'un système extérieur $\mathcal{S}$, décrit par $\mathcal{N}$ coordonnées généralisées $\mathcal{Q}_j$. Nous supposerons que les coordonnées généralisées sont bien séparées: l'état de S est complètement déterminé par les q_i seuls. Nous supposerons le système $\mathcal{S}$ suffisamment "gros" pour que l'interaction avec S ait une influence négligeable sur sa dynamique (c'est bien sûr le cas pour tout mouvement réaliste dans le champ de pesanteur terrestre). Nous pouvons alors considérer les $\mathcal{Q}_j$ comme des données du problème.

La fonction de Lagrange du système $S + \mathcal{S}$ s'écrit $L = T_S + T_{\mathcal{S}} - U(q_i; \mathcal{Q}_i)$, où U est l'énergie potentielle dont dérivent toutes les forces du problème (y compris l'interaction entre S et $\mathcal{S}$), T_S et $T_{\mathcal{S}}$ les énergies cinétiques des deux systèmes. Comme nous avons supposé que les deux ensembles de coordonnées généralisées sont bien séparés, T_S n'est fonction que des q_i et de leurs dérivées et $T_{\mathcal{S}}$ n'est fonction que des seuls $\mathcal{Q}_j$ et de leurs dérivées. $T_{\mathcal{S}}$ peut donc être considérée comme une fonction donnée du temps. Une telle fonction étant aussi une dérivée totale par rapport au temps, elle ne joue aucun rôle dans les équations de Lagrange et peut être supprimée. De la même manière, U peut être écrite comme une fonction des q_i seuls et du temps (la dépendance en temps reflétant la dynamique des $\mathcal{Q}_j$). On a finalement: $L = T_S(q_i, \dot{q}_i) - U(q_i, t)$, lagrangien décrivant la dynamique du système S seul.

Ce type de problème étant très fréquemment rencontré en mécanique, nous allons l'illustrer par deux exemples.

Premier exemple

Cet exemple, particulièrement trivial, ne présente guère que l'intérêt d'appliquer les notions introduites dans les paragraphes précédents dans une situation où la mise en équations et les calculs ne présentent aucune difficulté technique. Nous considérerons donc la "machine d'Atwood", pont aux ânes des classes élémentaires (voir figure 1.3). Deux masses m_1 et m_2, astreintes à se déplacer verticalement dans le champ de pesanteur (accélération $\mathbf{g}$), pendent aux deux extrémités d'une ficelle passant sur une poulie (tout cela étant inextensible, sans frottements...). Une seule coordonnée généralisée suffit à décrire ce problème en tenant compte des liaisons (mouvement selon la verticale, longueur de ficelle constante). On prendra la position x de la masse m_1 le long d'un axe vertical descendant. La vitesse de m_1 est $\dot{x}$, celle de m_2, $-\dot{x}$. On a alors $T = (m_1 + m_2)\dot{x}^2/2$ et $U = -m_1 g x - m_2 g(\ell - x)$, où ℓ est une constante. A une constante additive près on a donc $U = (m_2 - m_1)gx$ et

$$L = \frac{1}{2}(m_1 + m_2)\dot{x}^2 - (m_2 - m_1)gx \ . \tag{1.13}$$

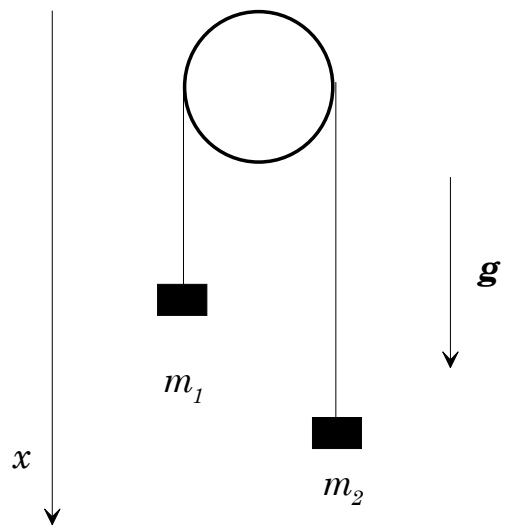

Figure 1.3: La "machine d'Atwood", exemple élémentaire de mouvements avec liaisons dans un champ extérieur. Deux masses différentes sont reliées par une corde inextensible passant sur une poulie. Le mouvement des masses est uniformément accéléré.

L'unique équation de Lagrange s'écrit alors trivialement:

$$\frac{d}{dt}\frac{\partial L}{\partial \dot{x}} = (m_1 + m_2)\ddot{x} = \frac{\partial L}{\partial x} = -(m_2 - m_1)g \ , \tag{1.14}$$

d'où on déduit bien évidemment un mouvement uniformément accéléré avec l'accélération $(m_1 - m_2)g/(m_1 + m_2)$

Pour un exemple aussi trivial, l'écriture du formalisme lagrangien n'apporte pas de simplification décisive. Il eut été aussi simple d'écrire, pour ce problème à un seul degré de liberté, le théorème de l'énergie cinétique. En revanche, l'écriture du principe fondamental de la dynamique nous aurait contraint à faire intervenir des inconnues supplémentaires: les tensions des fils. La puissance de l'écriture lagrangienne ne peut s'exprimer que sur des problèmes un peu plus complexes.

Deuxième exemple

Nous considérerons dans ce paragraphe l'exemple du double pendule (voir figure 1.1). Notons tout de suite que la mise en équations de ce problème est très pénible par les techniques standard de mécanique. Il faut en effet faire intervenir la tension des fils, dont on ne se débarrasse qu'au prix de manipulations fastidieuses. Notons aussi que, pour ce problème à deux degrés de liberté, le théorème de l'énergie cinétique ne nous est d'aucune utilité.

Nous allons écrire les équations de Lagrange. Les deux coordonnées généralisées q_1 et q_2 coïncident avec les angle θ_1 et θ_2 que les pendules font avec la verticale. L'écriture des positions et vitesses des deux masses en fonction des coordonnées généralisées et des longueurs l_1 et l_2 des deux pendules, ne présente aucune difficulté. On en déduit les expressions des énergies cinétiques et potentielles:

$$\begin{aligned}
T &= T_1 + T_2 \\
T_1 &= \frac{1}{2}m_1 l_1^2 \dot{\theta}_1^2 && (1.15)\\
T_2 &= \frac{1}{2}m_2 \left[l_1^2\dot{\theta}_1^2 + l_2^2\dot{\theta}_2^2 + 2l_1 l_2 \cos(\theta_1 - \theta_2)\dot{\theta}_1\dot{\theta}_2 \right] && (1.16)\\
U &= U_1 + U_2 \\
U_1 &= -m_1 g l_1 \cos(\theta_1) && (1.17)\\
U_2 &= -m_2 g l_2 \cos(\theta_2) - m_2 g l_1 \cos(\theta_1) \ , && (1.18)
\end{aligned}$$

d'où on tire l'expression de la fonction de Lagrange L. L'écriture des deux équations de Lagrange ne pose alors d'autre difficulté qu'algébrique. La mise en équations de ce problème, pratiquement impossible en appliquant simplement le principe fondamental de la dynamique, ne présente aucune difficulté avec l'approche lagrangienne.

La résolution explicite du système différentiel obtenu est toutefois impossible. Il s'agit en effet d'équations non linéaires couplées (à cause des termes en $\cos\theta$). On ne peut obtenir de solutions explicites, comme dans tous les problèmes de pendules, que dans le cadre d'une approximation linéaire, valide dans le domaine des petites oscillations. Nous donnons ici seulement les étapes essentielles du calcul, dont les détails seront facilement retrouvés par le lecteur. Les équations de Lagrange pour θ_1 et θ_2 s'écrivent respectivement après linéarisation:

$$(m_1+m_2)l_1\ddot{\theta}_1 + m_2 l_2\ddot{\theta}_2 + (m_1+m_2)g\theta_1 = 0 \tag{1.19}$$

$$l_2\ddot{\theta}_2 + l_1\ddot{\theta}_1 + g\theta_2 = 0\ . \tag{1.20}$$

Un tel système de deux équations linéaires à coefficients constants se résout par les méthodes standard. Les solutions sont des combinaisons linéaires de solutions en $\exp(-i\omega t)$. Les valeurs possibles de ω sont celles qui annulent le déterminant caractéristique:

$$\begin{vmatrix} (m_1+m_2)(g-l_1\omega^2) & -m_2 l_2\omega^2 \\ -l_1\omega^2 & g-l_2\omega^2 \end{vmatrix}\ . \tag{1.21}$$

On obtient alors une équation bicarrée en ω. Les solutions s'écrivent $\pm\omega_1$ et $\pm\omega_2$. La solution générale est alors une superposition de deux mouvements oscillatoires aux fréquences ω_1 et ω_2, de phase et d'amplitudes arbitraires. L'écriture explicite de ces phases et amplitudes en fonction des conditions initiales est pénible, mais sans difficultés.

La recherche des deux fréquences ω_1 et ω_2 est un exemple particulièrement simple de recherches de "modes propres" dans des systèmes d'oscillateurs couplés. Très généralement, un système d'oscillateurs linéaires couplés possédant p degrés de liberté, admet p fréquences propres. Chacune de ces fréquences correspond à une configuration particulière des mouvements des degrés de liberté, conduisant à un mouvement purement harmonique (on pensera par exemple aux modes d'oscillations symétriques et antisymétriques de deux pendules couplés par un ressort). Dans le cas présent, il est instructif d'examiner les comportements asymptotiques des deux modes quand $m_1 \to \infty$ ou $m_2 \to 0$, ce que nous laissons au lecteur.

1.3.4 Lagrangien de particules chargées dans un champ

Nous abordons ici un exemple très important d'écriture de fonction de Lagrange, d'abord en raison de son utilité pratique, ensuite en raison de l'usage que nous en ferons dans la suite du cours d'électromagnétisme. Nous utiliserons une démarche très pragmatique, en cherchant la forme la plus simple de la fonction de Lagrange qui redonne la force de Lorentz.

Nous considérerons donc un ensemble de N particules, que nous supposerons décrites par leurs cordonnées cartésiennes standard (la généralisation à d'autres systèmes de coordonnées ne posant que des problèmes algébriques) $\mathbf{r}_\alpha$. Ces particules, chargées, sont placées dans un champ électrique $\mathbf{E}$ et un champ magnétique $\mathbf{B}$, imposés de l'extérieur. Pour alléger les notations, nous poserons $\mathbf{E}_\alpha = \mathbf{E}(\mathbf{r}_\alpha, t)$ et $\mathbf{B}_\alpha = \mathbf{B}(\mathbf{r}_\alpha, t)$. Nous introduirons enfin les potentiels vecteur $\mathbf{A}_\alpha = \mathbf{A}(\mathbf{r}_\alpha, t)$ et scalaire $V_\alpha = V(\mathbf{r}_\alpha, t)$ permettant de calculer les champs par les relations: $\mathbf{E}_\alpha = -\partial\mathbf{A}_\alpha/\partial t - \boldsymbol{\nabla} V_\alpha$ et $\mathbf{B}_\alpha = \boldsymbol{\nabla}\times\mathbf{A}_\alpha$, où les dérivées spatiales sont à prendre par rapport à la position de la particule α. Notons que la dérivée temporelle du potentiel vecteur est une dérivée partielle. Le champ électromoteur ne fait en effet intervenir que la dépendance explicite en temps du champ magnétique.

Le principe fondamental de la dynamique pour la particule α s'écrit alors:

$$m\ddot{\mathbf{r}}_\alpha = \mathbf{F}_\alpha = q_\alpha(\mathbf{E}_\alpha + \mathbf{v}_\alpha\times\mathbf{B}_\alpha)\ . \tag{1.22}$$

Nous allons transformer cette équation pour la mettre sous une forme qui rappelle celle d'une équation de Lagrange, ce qui nous permettra d'intuiter une forme simple pour la fonction de Lagrange de ce problème.

En faisant intervenir l'expression des champs en fonction des potentiels, on peut écrire:

$$\mathbf{F}_\alpha = q_\alpha \left(-\frac{\partial \mathbf{A}_\alpha}{\partial t} + \mathbf{v}_\alpha \times (\boldsymbol{\nabla} \times \mathbf{A}_\alpha) - \boldsymbol{\nabla} V_\alpha \right) \ . \tag{1.23}$$

En place de la dérivée partielle par rapport au temps de $\mathbf{A}_\alpha$, on souhaite faire apparaître une dérivée totale par rapport au temps (rappelons que, dans les équations de Lagrange, les dérivées par rapport au temps sont totales). La variation de $\mathbf{A}_\alpha$ par rapport au temps provient de deux causes. D'abord, il peut exister une variation explicite du potentiel vecteur par rapport au temps (exprimée ci–dessus par la dérivée partielle par rapport au temps). Même en l'absence de cette dépendance explicite en temps, le potentiel vecteur "vu" par a particule α varie en raison du mouvement de la particule dans le champ inhomogène. La variation correspondante peut s'écrire simplement $(\mathbf{v}_\alpha \cdot \boldsymbol{\nabla})\mathbf{A}_\alpha$, terme parfois nommé "dérivée hydrodynamique". On a alors:

$$\frac{d\mathbf{A}_\alpha}{dt} = \frac{\partial \mathbf{A}_\alpha}{\partial t} + (\mathbf{v}_\alpha \cdot \boldsymbol{\nabla})\mathbf{A}_\alpha \ , \tag{1.24}$$

et donc

$$\mathbf{F}_\alpha = q_\alpha \left(-\frac{d\mathbf{A}_\alpha}{dt} + (\mathbf{v}_\alpha \cdot \boldsymbol{\nabla})\mathbf{A}_\alpha + \mathbf{v}_\alpha \times (\boldsymbol{\nabla} \times \mathbf{A}_\alpha) - \boldsymbol{\nabla} V_\alpha \right) \ . \tag{1.25}$$

On reconnaît dans les deux termes centraux deux des termes du développement de $\boldsymbol{\nabla}_{\mathbf{r}_\alpha}(\mathbf{v}_\alpha \cdot \mathbf{A}_\alpha)$. Les deux autre termes de ce développement qui font intervenir des dérivées partielles de $\mathbf{v}_\alpha$ par rapport à la position de la particule α sont manifestement nuls. On obtient donc finalement:

$$\mathbf{F}_\alpha = q_\alpha \left(-\frac{d\mathbf{A}_\alpha}{dt} - \boldsymbol{\nabla} \left[V_\alpha - \mathbf{A}_\alpha \cdot \mathbf{v}_\alpha \right] \right) \ . \tag{1.26}$$

En réécrivant alors le premier terme de (1.22) sous la forme:

$$\frac{d}{dt} \boldsymbol{\nabla}_{\mathbf{v}_\alpha} \frac{1}{2} m_\alpha \mathbf{v}_\alpha^2 \ , \tag{1.27}$$

et en remarquant que

$$\frac{d\mathbf{A}_\alpha}{dt} = \frac{d}{dt} \boldsymbol{\nabla}_{\mathbf{v}_\alpha} \mathbf{A}_\alpha \cdot \mathbf{v}_\alpha \tag{1.28}$$

on peut mettre le principe fondamental sous la forme:

$$\begin{aligned} \frac{d}{dt} \boldsymbol{\nabla}_{\mathbf{v}_\alpha} \left[\frac{1}{2} m_\alpha \mathbf{v}_\alpha^2 + q_\alpha \mathbf{A}_\alpha \cdot \mathbf{v}_\alpha \right] &= -q_\alpha \boldsymbol{\nabla}_{\mathbf{r}_\alpha} \left[V_\alpha - \mathbf{A}_\alpha \cdot \mathbf{v}_\alpha \right] \\ &= \boldsymbol{\nabla}_{\mathbf{r}_\alpha} \left[\frac{1}{2} m_\alpha \mathbf{v}_\alpha^2 - q_\alpha (V_\alpha - \mathbf{A}_\alpha \cdot \mathbf{v}_\alpha) \right] \ , \end{aligned} \tag{1.29}$$

le gradient par rapport à la position de l'énergie cinétique étant évidemment nul. De la même manière, on peut ajouter dans le gradient du membre de gauche un terme proportionnel au potentiel, $q_\alpha V_\alpha$, qui ne dépend manifestement pas des vitesses. On obtient alors la forme de l'équation de Lagrange en coordonnées cartésiennes, à condition de poser:

$$L = \sum_\alpha \left[T_\alpha - q_\alpha V_\alpha + q_\alpha \mathbf{A}_\alpha \cdot \mathbf{v}_\alpha \right] \ . \tag{1.30}$$

Nous admettrons que cette forme du lagrangien est effectivement convenable. Quelques remarques s'imposent à ce point:

- Le terme "d'énergie potentielle" que nous obtenons ici dépend explicitement de la vitesse des particules. Nous retrouvons ici, sous une autre forme, la non conservativité des forces électromagnétiques: l'énergie mécanique totale n'est pas constante dans le cas le plus général. Ce n'est que si le potentiel vecteur est indépendant du temps (au sens d'une dérivée partielle par rapport au temps) et de plus uniforme que le dernier terme du lagrangien se ramène à une dérivée totale (celle de $\mathbf{A}_\alpha \cdot \mathbf{r}_\alpha$), qui peut être oubliée. Le potentiel électrostatique joue alors le rôle d'une vraie énergie potentielle, ne dépendant que des positions des particules. Ce n'est donc que dans le cadre de l'électrostatique que les forces électromagnétiques sont conservatives, un résultat bien connu sur lequel nous jetons ici un éclairage nouveau. Notons que la non conservation de l'énergie mécanique ne signifie pas une violation de la conservation de l'énergie totale. Nous verrons en effet en relativité que l'énergie et l'impulsion totales sont toujours conservées, à condition de faire intervenir le champ dans le bilan.

- Encore une fois, les champs sont ici imposés de l'extérieur, et ne font pas partie des variables dynamiques du problème. Pour traiter le problème complet de particules en interaction avec leur propre champ, il faudrait adjoindre aux équations de Lagrange que nous venons d'écrire les équations de Maxwell permettant de calculer les champs en fonction des mouvements des particules. Ce n'est que dans la partie suivante que nous apprendrons à le faire de manière consistante dans un formalisme lagrangien englobant le champ.

- L'expression de notre lagrangien fait explicitement intervenir les potentiels. Il nous faut donc vérifier que les équations de Lagrange ne sont pas modifiées quand on effectue une transformation de jauge sur les potentiels. Rappelons en effet que les champs "physiques" ne sont pas modifiés si on effectue la transformation de "jauge":

$$\mathbf{A} \longrightarrow \mathbf{A}' = \mathbf{A} + \boldsymbol{\nabla}\phi \tag{1.31}$$

$$V \longrightarrow V' = V - \frac{\partial \phi}{\partial t} , \tag{1.32}$$

où ϕ est une fonction arbitraire de l'espace et du temps. Le nouveau lagrangien L' s'exprime alors facilement en fonction de l'ancien, en notant $\phi_\alpha = \phi(\mathbf{r}_\alpha, t)$:

$$\begin{aligned} L' &= L + \sum_\alpha q_\alpha \left(\frac{\partial \phi_\alpha}{\partial t} + (\boldsymbol{\nabla}\phi_\alpha) \cdot \mathbf{v}_\alpha \right) \\ &= L + \sum_\alpha q_\alpha \left(\frac{\partial \phi_\alpha}{\partial t} + (\mathbf{v}_\alpha \cdot \boldsymbol{\nabla})\phi_\alpha \right) \\ &= L + \sum_\alpha q_\alpha \frac{d\phi_\alpha}{dt} , \end{aligned} \tag{1.33}$$

en reconnaissant dans les deux derniers termes de l'équation centrale la dérivée totale par rapport au temps de ϕ_α, somme de la dérivée partielle et de la "dérivée hydrodynamique". L' et L ne diffèrent donc que par une dérivée totale par rapport au temps qui ne modifie pas le contenu des équations de Lagrange, comme nous nous y attendions.

1.4 Généralisations

Nous traiterons dans ce paragraphe de deux extensions du formalisme lagrangien[3]. Nous traiterons d'abord le cas de forces (non électromagnétiques) ne dérivant pas d'un potentiel. Nous verrons par exemple comment traiter le cas des forces de frottement dans le cadre d'un principe variationnel. Dans

[3]Les résultats de ce chapitre ne seront pratiquement pas utilisés dans la suite du cours. On pourra donc le sauter en première lecture

la seconde partie de ce paragraphe, nous nous pencherons sur le cas des systèmes où les mouvements sont sujets à des liaisons non holonomes. Nous verrons comment on peut facilement incorporer de telles liaisons dans le formalisme lagrangien.

1.4.1 Forces ne dérivant pas d'une énergie potentielle

On suppose connues les forces $\mathbf{F}_\alpha$ s'exerçant sur la particule α. Notre problème est de trouver une expression variationnelle de la dynamique et d'en déduire les équations de Lagrange correspondantes, même si les forces ne dérivent pas d'une énergie potentielle.

Revenons d'abord au cas des forces conservatives, pour mettre le principe de moindre action sous une forme un peu différente. Tel que nous l'avons écrit, le principe de moindre action stipule que $\delta S = \delta \int_{t_1}^{t_2} L\, dt = 0$, si on estime la variation de l'action entre la trajectoire effectivement suivie et une autre trajectoire infiniment voisine compatible avec les conditions aux limites. On peut aussi écrire $\int_{t_1}^{t_2} (\delta T - \delta U)\, dt = 0$. Si on note $\delta \mathbf{r}_\alpha$ l'écart pour la particule α entre la trajectoire normale et la trajectoire variée, on a $\delta U = -\sum_\alpha \mathbf{F}_\alpha \cdot \delta \mathbf{r}_\alpha$, le principe de moindre action s'écrivant alors

$$\int_{t_1}^{t_2} (\delta T + \sum_\alpha \mathbf{F}_\alpha \cdot \delta \mathbf{r}_\alpha)\, dt = 0 \ . \tag{1.34}$$

On peut exprimer les $\delta \mathbf{r}_\alpha$ en fonction des accroissements δq_i des coordonnées généralisées:

$$\sum_\alpha \mathbf{F}_\alpha \cdot \delta \mathbf{r}_\alpha = \sum_\alpha \sum_i \mathbf{F}_\alpha \cdot \frac{\partial \mathbf{r}_\alpha}{\partial q_i} \delta q_i \ . \tag{1.35}$$

Posons alors

$$Q_i = \sum_\alpha \mathbf{F}_\alpha \cdot \frac{\partial \mathbf{r}_\alpha}{\partial q_i} \ . \tag{1.36}$$

Nous désignerons Q_i comme la force généralisée correspondant à la coordonnée généralisée q_i[4]. Avec cette notation, le principe de moindre action s'écrit:

$$\int_{t_1}^{t_2} (\delta T + \sum_i Q_i \delta q_i)\, dt = 0 \ . \tag{1.37}$$

Nous avons établi cette formulation en supposant que les forces dérivaient d'une énergie potentielle. Nous admettrons qu'elle reste valable même si ce n'est pas le cas. Nous allons maintenant dériver de cette expression du principe de moindre action les équations de Lagrange correspondantes. L'approche est très similaire à celle que nous employâmes dans le paragraphe 1.2. T étant une fonction des q_i de leurs dérivées et du temps, on peut écrire

$$\delta T = \frac{\partial T}{\partial q_i} \delta q_i + \frac{\partial T}{\partial \dot{q}_i} \delta \dot{q}_i \ . \tag{1.38}$$

Une intégration par parties élémentaire donne alors:

$$\int_{t_1}^{t_2} \frac{\partial T}{\partial \dot{q}_i} \delta \dot{q}_i\, dt = \left[\frac{\partial T}{\partial \dot{q}_i} \delta q_i \right]_1^2 - \int_{t_1}^{t_2} \frac{d}{dt} \frac{\partial T}{\partial \dot{q}_i} \delta q_i\, dt \ . \tag{1.39}$$

Le terme tout intégré est identiquement nul, puisque les deux trajectoires coïncident aux limites. En reportant cette expression dans le principe de moindre action et en écrivant que l'intégrale doit s'annuler quelles que soient les variations des coordonnées indépendantes q_i, on obtient les équations de Lagrange:

$$\frac{d}{dt} \frac{\partial T}{\partial \dot{q}_i} - \frac{\partial T}{\partial q_i} = Q_i \ . \tag{1.40}$$

[4] "Force" qui n'a pas nécessairement la dimension d'un force et qui ne peut être en général attribuée à une particule particulière.

L'écriture de ce système d'équations et sa résolution ne présente aucune difficulté si on connaît les forces s'exerçant sur les particules.

Dans le cas particulier où la force généralisée Q_i peut se mettre sous la forme:

$$Q_i = -\frac{\partial U}{\partial q_i} + \frac{d}{dt}\frac{\partial U}{\partial \dot{q}_i} \qquad \text{où} \quad U = U(q_i, \dot{q}_i, t) \ , \tag{1.41}$$

l'équation de Lagrange (1.40) se met sous la forme standard avec $L = T - U$. On retrouve ainsi le cas des forces conservatives mais aussi le cas des forces électromagnétiques. Notons qu'il n'est pas toujours possible de mettre la force généralisée sous cette forme (nous verrons que c'est le cas dans l'exemple traité à la fin de ce paragraphe). Soulignons aussi que nous permettons à U de dépendre des vitesses. Il est donc possible (comme nous l'avons vu pour les forces électromagnétiques) d'englober certaines forces non conservatives dans des équations de Lagrange sous la forme standard.

Très souvent, une partie des forces dérive d'un potentiel (au moins au sens généralisé exprimé par l'équation 1.41). Dans ce cas, en appelant U "l'énergie potentielle" associée aux forces en dérivant et Q_i les forces généralisées correspondant aux forces ne dérivant pas de U, on écrira les équations de Lagrange sous la forme:

$$\frac{d}{dt}\frac{\partial L}{\partial \dot{q}_i} - \frac{\partial L}{\partial q_i} = Q_i \qquad \text{avec} \quad L = T - U. \tag{1.42}$$

Nous appliquerons, à titre d'exemple, cette démarche à un système où s'exercent des forces de frottement proportionnelles à la vitesse. De telles forces ne sont manifestement pas conservatives. Nous poserons donc (en oubliant pour l'instant toutes les autres forces) $\mathbf{F}_\alpha = -k_\alpha \mathbf{v}_\alpha$. Cette force peut s'écrire $\mathbf{F}_\alpha = -\boldsymbol{\nabla}_{\mathbf{v}_\alpha}\mathcal{F}$ à condition de poser

$$\mathcal{F} = \frac{1}{2}\sum_\alpha k_\alpha \mathbf{v}_\alpha^2 \ . \tag{1.43}$$

Les forces généralisées Q_i s'écrivent:

$$Q_i = -\sum_\alpha \boldsymbol{\nabla}_{\mathbf{v}_\alpha}\mathcal{F} \cdot \frac{\partial \mathbf{r}_\alpha}{\partial q_i} \ . \tag{1.44}$$

En remarquant alors que:

$$\frac{\partial \dot{\mathbf{r}}_\alpha}{\partial \dot{q}_i} = \frac{\partial \mathbf{r}_\alpha}{\partial q_i} \ , \tag{1.45}$$

on a

$$Q_i = -\sum_\alpha \boldsymbol{\nabla}_{\mathbf{v}_\alpha}\mathcal{F} \cdot \frac{\partial \mathbf{v}_\alpha}{\partial \dot{q}_i} = -\frac{\partial \mathcal{F}}{\partial \dot{q}_i} \ . \tag{1.46}$$

On voit bien qu'une telle expression n'est pas compatible en général avec (1.41).

Les équations de Lagrange en présence de frottements visqueux s'écrivent finalement:

$$\frac{d}{dt}\frac{\partial L}{\partial \dot{q}_i} - \frac{\partial L}{\partial q_i} + \frac{\partial \mathcal{F}}{\partial \dot{q}_i} = 0 \ , \tag{1.47}$$

où L tient compte des éventuelles forces "conservatives".

Pour un mouvement unidimensionnel d'une particule soumise uniquement à une force de frottement, on a $\mathcal{F} = k\dot{x}^2/2$, et $L = m\dot{x}^2/2$. L'équation de Lagrange s'écrit donc $m\ddot{x} + k\dot{x} = 0$, ce qui coïncide bien avec l'expression du principe fondamental (encore une fois, le formalisme lagrangien n'apporte rien pour un problème aussi élémentaire).

1.4.2 Cas des liaisons non holonomes

Nous traiterons ici le cas de systèmes où les variables dynamiques peuvent être reliées par des liaisons non holonomes, c'est à dire ne se mettant pas simplement sous la forme de relations entre les coordonnées cartésiennes (éventuellement dépendant du temps). Nous nous contenterons de traiter le cas où les contraintes s'expriment par des relations entre les vitesses des différentes particules composant le système. Ce cas permet en effet de traiter de la plupart des liaisons du type "roulement sans glissement" dont l'importance est considérable en mécanique du solide. Les techniques que nous introduirons ici peuvent être étendues à d'autres types de contraintes.

Nous procéderons en deux étapes. Nous ne tiendrons d'abord compte que des liaisons holonomes, en introduisant des coordonnées généralisées qui seraient indépendantes si ces liaisons étaient les seules. Nous reviendrons sur le principe de moindre action pour montrer que l'écriture des équations de Lagrange en termes de ces variables dépendantes est impossible. Nous verrons ensuite qu'en introduisant des variables supplémentaires, les multiplicateurs de Lagrange, on peut obtenir un système d'équations différentielles indépendantes permettant, au moins en principe, de résoudre le problème. Nous interpréterons physiquement, enfin, ces variables supplémentaires et montrerons qu'elles décrivent les forces généralisées associées aux liaisons non holonomes.

En tenant compte uniquement des liaisons holonomes, nous définissons un ensemble de n coordonnées généralisées q_i. Pour fixer les idées, nous supposerons que toutes les forces (à part celles associées aux liaisons) dérivent d'une énergie potentielle, éventuellement généralisée comme au paragraphe précédent. Les liaisons non holonomes font que ces n coordonnées généralisées ne sont pas indépendantes. Nous supposerons qu'il existe m relations entre elles (m liaisons) et nous supposerons que ces relations relient les dérivées des q_i par rapport au temps et peuvent donc s'écrire:

$$\sum_{k=1}^{n} a_{lk} dq_k + a_{lt} dt = 0, \qquad (l = 1...m) \ , \tag{1.48}$$

où les coefficients $a_{lk,t}$ sont indépendants du temps.

Reprenons maintenant la dérivation des équations de Lagrange à partir du principe de moindre action, comme dans le paragraphe 1.2. La comparaison de l'action sur la trajectoire effectivement suivie et sur la trajectoire variée n'est pas modifiée, et on écrit:

$$\delta S = \int_{t_1}^{t_2} \sum_k \left[\frac{\partial L}{\partial q_k} - \frac{d}{dt} \frac{\partial L}{\partial \dot{q}_k} \right] \delta q_k \, dt \ . \tag{1.49}$$

De cette équation, on ne peut déduire des équations de Lagrange pour les q_k. Ce ne sont pas en effet des variables indépendantes, dont les accroissements puissent être choisis arbitrairement. Les δq_k sont en fait reliés par les m relations:

$$\sum_k a_{lk} \delta \dot{q}_k = 0 \ . \tag{1.50}$$

Elles résultent simplement de (1.48) en constatant que les a_{lt} s'éliminent quant on écrit les liaisons entre les accroissements des dérivées entre trajectoire de référence et trajectoire variée. Par intégration, les sommes $\sum_k a_{lk} \delta q_k$ sont constantes. Les accroissements de trajectoire s'annulant aux instants initiaux et finaux, cette constante est nulle et on a finalement:

$$\sum_k a_{lk} \delta q_k = 0 \ . \tag{1.51}$$

En d'autres termes, nous avons a priori trop d'équations pour les $n-m$ variables vraiment indépendantes. Plutôt que d'éliminer des variables superfétatoires, nous allons introduire m inconnues supplémentaires (dont nous verrons qu'elles ont un sens physique). Nous considérons donc m fonctions supplémentaires du temps, pour l'instant arbitraires, $\lambda_l(t)$. Elles vérifient bien sûr:

$$\lambda_l \sum_k a_{lk} \delta q_k = 0 \ , \tag{1.52}$$

et donc

$$\int_{t_1}^{t_2} \sum_{k,l} \lambda_l a_{lk} \delta q_k = 0 \ . \tag{1.53}$$

En introduisant cette relation dans l'expression du principe variationnel, on le met sous la forme:

$$\delta S = \int_{t_1}^{t_2} \sum_{k=1}^{n} \left[\frac{\partial L}{\partial q_k} - \frac{d}{dt} \frac{\partial L}{\partial \dot{q}_k} + \sum_{l=1}^{m} \lambda_l a_{lk} \right] \delta q_k \, dt \ . \tag{1.54}$$

Prenons comme variables indépendantes les $n - m$ premières coordonnées généralisées. Au moins formellement, les m dernières peuvent être calculées en fonction de celles ci. Nous choisissons alors les m fonctions λ_l de telle manière que:

$$\frac{\partial L}{\partial q_k} - \frac{d}{dt} \frac{\partial L}{\partial \dot{q}_k} + \sum_{l=1}^{m} \lambda_l a_{lk} = 0 \ , \tag{1.55}$$

pour toutes les valeurs de k entre $n - m + 1$ et n. Ce choix est toujours possible. Les λ_l sont en effet définis comme solutions d'un système linéaire dont le déterminant, formé des coefficients a_{lk} est non nul si les m relations exprimant les contraintes non holonomes sont linéairement indépendantes.

Le principe variationnel s'écrit alors:

$$\delta S = \int_{t_1}^{t_2} \sum_{k=1}^{m} \left[\frac{\partial L}{\partial q_k} - \frac{d}{dt} \frac{\partial L}{\partial \dot{q}_k} + \sum_{l=1}^{m} \lambda_l a_{lk} \right] \delta q_k \, dt \qquad (k = 1...m) \ . \tag{1.56}$$

Les m premières coordonnées et leurs accroissements étant indépendants, cette intégrale n'est nulle que si les m quantités entre crochets sont identiquement nulles sur la trajectoire effectivement suivie. En ajoutant à ces m équations de Lagrange les $n - m$ relations (1.55) et les m contraintes (1.48), on obtient enfin un système de $n + m$ équations à $n + m$ inconnues (q_i et λ_l):

$$\left\{ \begin{array}{ll} \dfrac{\partial L}{\partial q_k} - \dfrac{d}{dt} \dfrac{\partial L}{\partial \dot{q}_k} + \displaystyle\sum_{l=1}^{m} \lambda_l a_{lk} = 0 \ , & (k = 1 \ldots n) \\ \displaystyle\sum_{k=1}^{m} a_{lk} \dot{q}_k + a_{lt} = 0 \ , & (l = 1 \ldots m) \ . \end{array} \right. \ , \tag{1.57}$$

dont la résolution donne la trajectoire effectivement suivie.

Nous avons pu tenir compte des liaisons non holonomes supplémentaires en introduisant des variables additionnelles, ce qui n'est guère économique, surtout quand il s'agit de résoudre effectivement le système. Cependant, les λ_l possèdent une interprétation physique qui rend leur obtention importante. Nous pourrions en effet "simuler" l'effet des liaisons non holonomes en appliquant des forces supplémentaires dans le système (tout mouvement peut toujours être vu comme résultant de forces et pas de contraintes). En fait, ces forces ont une réalité physique. Elles correspondent, par exemple, aux frottements responsables d'un "roulement sans glissement". Ces forces, ne dérivant pas en général d'un potentiel, doivent intervenir dans les équations de Lagrange sous la forme de "forces généralisées" Q_k. Les n équations de Lagrange de ce système sans contraintes supplémentaires s'écriraient, avec la même fonction de Lagrange:

$$\frac{\partial L}{\partial q_k} - \frac{d}{dt} \frac{\partial L}{\partial \dot{q}_k} + Q_k = 0 \ , \qquad (k = 1 \cdots n) \ . \tag{1.58}$$

En comparant ces équations avec celles de (1.57), on voit que les Q_k doivent être définis par:

$$Q_k = \sum_{l=1}^{m} \lambda_l a_{lk} \ . \tag{1.59}$$

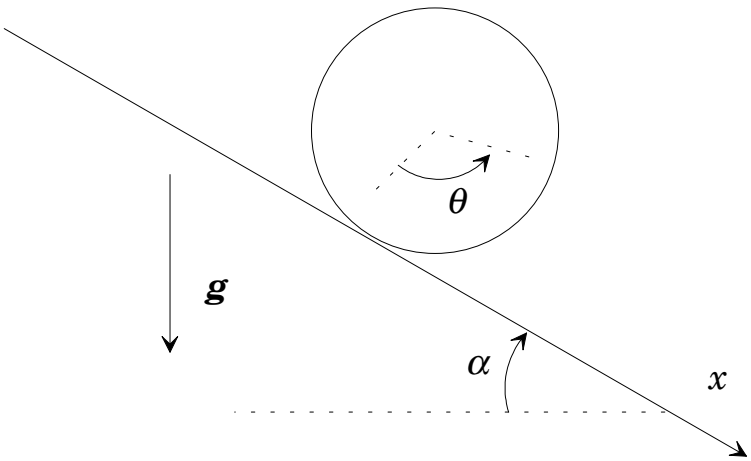

Figure 1.4: Roulement sans glissement d'un cylindre sur un plan incliné. Le mouvement est décrit a priori par deux coordonnées généralisées: la position du point de contact le long du plan et l'angle de rotation du cylindre par rapport à une référence arbitraire. Dans le cas d'un roulement sans glissement, ces deux coordonnées sont reliées par une relation non holonome.

Les fonctions supplémentaires λ_l ne sont donc que des combinaisons linéaires des forces généralisées correspondant aux contraintes. Leur obtention à partir de (1.57) permet donc de calculer les forces de liaison, ce qui justifie amplement leur intérêt.

Pour illustrer ce paragraphe, nous considérerons un problème trivial de dynamique des solides: le roulement sans glissement d'un cylindre creux (la masse est entièrement distribuée sur la surface extérieure) sur un plan incliné[5] (voir figure 1.4). Si on ne tient compte que des liaisons holonomes (le cylindre est posé sur le plan, il ne se déplace que dans la direction x...), nous pouvons décrire le mouvement par deux coordonnées généralisées. $q_1 = x$ est simplement l'abscisse du point de contact sur le plan, $q_2 = \theta$ l'angle de rotation du cylindre autour de son axe. Le roulement sans glissement impose l'unique contrainte non holonome supplémentaire $R\dot{\theta} + \dot{x} = 0$, où R est le rayon du cylindre. Nous n'avons donc que deux coefficients a_{lk}: $a_{11} = 1$ et $a_{12} = R$.

La fonction de Lagrange est $L = T - U$. L'énergie cinétique s'écrit:

$$T = \frac{1}{2}M\dot{q}_1^2 + \frac{1}{2}MR^2\dot{q}_2^2 \ , \tag{1.60}$$

où M est la masse du cylindre, et MR^2 son moment d'inertie par rapport à son axe. On a aussi, dans le champ de pesanteur, $U = -Mgq_1 \sin\alpha$. Avec une seule liaison, il existe un seul multiplicateur de Lagrange λ_1. Les deux équations de Lagrange et la contrainte nous fournissent alors un système de 3 équations à 3 inconnues:

$$\begin{cases} \dfrac{\partial L}{\partial q_1} - \dfrac{d}{dt}\dfrac{\partial L}{\partial \dot{q}_1} + \lambda_1 = 0 \\ \dfrac{\partial L}{\partial q_2} - \dfrac{d}{dt}\dfrac{\partial L}{\partial \dot{q}_2} - R\lambda_1 = 0 \\ \dot{q}_1 - R\dot{q}_2 = 0 \end{cases} \ . \tag{1.61}$$

[5]Ce problème est en fait si élémentaire qu'il peut être traité en quelques lignes par le théorème de l'énergie cinétique.

On en déduit immédiatement:

$$\begin{cases} M\ddot{q}_1 - Mg\sin\alpha - \lambda_1 = 0 \\ \\ MR^2\ddot{q}_2 - R\lambda_1 = 0 \\ \\ \dot{q}_1 + R\dot{q}_2 = 0 \end{cases} . \tag{1.62}$$

Après quelques manipulations algébriques sans intérêt, on obtient un mouvement uniformément accéléré pour x avec l'accélération $g\sin(\alpha)/2$, un résultat bien connu de dynamique standard, et $\lambda_1 = -Mg\sin(\alpha)/2$. La force généralisée correspondant à la contrainte de roulement, Q_1 lui est égale. Enfin, on a simplement $Q_1 = F$, où F est la valeur algébrique de la force de frottement s'exerçant, tangentiellement au cylindre, au point de contact[6].

1.5 Lagrangien et lois de conservation

Dans tout phénomène physique il existe des quantités conservées au cours de l'évolution. Ces quantités, aussi appelées en mécanique "intégrales premières du mouvement", jouent un rôle important en fournissant des renseignements sur la dynamique, même si la trajectoire n'est pas explicitement connue. Par exemple, on peut donner beaucoup de caractéristiques générales des collisions en écrivant la conservation de l'impulsion ou de l'énergie, sans même connaître de façon détaillée la loi régissant l'interaction entre les particules[7].

Parmi toutes les quantités conservées dans le mouvement, certaines sont triviales et d'autres sans interprétation physique directe. Certaines en revanche, comme l'énergie ou l'impulsion, sont directement reliées à des propriétés physiques très fondamentales. Comme nous le verrons, il existe en général une quantité conservée associée à chaque invariance du système dans une transformation (à chaque propriété de symétrie, au sens le plus large). Nous verrons par exemple que la conservation de l'impulsion découle directement de l'invariance par translation dans l'espace.

Avant d'entrer dans le détail de ces lois de conservation, prouvons d'abord l'existence de quantités conservées dans une évolution lagrangienne. Nous considérerons dans tout ce paragraphe un système exempt des complications et généralisations introduites au paragraphe précédent. Nous ne considérerons donc que des liaisons holonomes. Comme les équations de Lagrange du système pour les n coordonnées généralisées q_i sont du second ordre en temps, la solution explicite du problème fait intervenir $2n$ constantes représentant les conditions aux limites (valeurs des coordonnées aux deux extrémités de la trajectoire). L'une de ces constantes peut toujours être mise sous la forme d'une origine arbitraire t_0 sur le temps. La solution la plus générale peut donc s'écrire:

$$q_i(t + t_0, C_1, \cdots, C_{2n-1}) \ , \dot{q}_i(t + t_0, C_1, \cdots, C_{2n-1}) \ , \tag{1.63}$$

où t_0 et les C_k sont les $2n$ constantes. Le mouvement étant déterminé, on peut inverser ces relations entre les $2n$ constantes et les cordonnées, et écrire:

$$t_0(q_i, \dot{q}_i, t) \ , C_k(q_i, \dot{q}_i, t) \ . \tag{1.64}$$

L'origine du temps joue un rôle particulier, et n'est pas à proprement parler une intégrale première du mouvement. En revanche, les $2n - 1$ C_k sont bien des fonctions de l'état dynamique du système qui restent constantes au cours du mouvement. Il y a donc, de façon très générale, au moins $2n - 1$ intégrales premières indépendantes dans un mouvement à n degrés de liberté.

Dans la suite de ce paragraphe, nous allons nous concentrer sur trois d'entre elles: l'impulsion, l'énergie et le moment cinétique et montrer comment elles se déduisent de propriétés de symétrie du lagrangien.

[6] Ce résultat peut être établi immédiatement à partir de l'accélération du mouvement

[7] On consultera avec profit le Landau de Mécanique sur ce problème que nous n'aborderons pas du tout dans ce cours.

1.5.1 Invariance par translation dans le temps: énergie

Nous considérons ici un système isolé. La première conséquence est que la fonction de Lagrange ne peut dépendre explicitement du temps. Il n'y a rien en effet dans l'environnement du système pour fixer une origine de temps. Nous supposerons de plus que les équations dynamiques de ce système sont des équations de Lagrange standard (sans forces généralisées). On exclut donc explicitement le cas où le système serait soumis à des forces de frottement (difficilement compatibles avec l'hypothèse d'un système isolé).

On a donc $\partial L/\partial t = 0$. La dérivée totale de L par rapport au temps, somme de sa dérivée partielle et des variations temporelles provenant de la variation des coordonnées généralisées, peut donc s'écrire:

$$\frac{dL}{dt} = \sum_i \frac{\partial L}{\partial \dot{q}_i}\ddot{q}_i + \sum_i \frac{\partial L}{\partial q_i}\dot{q}_i \ . \tag{1.65}$$

En utilisant les équations de Lagrange vérifiées par L et les q_i, on met cette dérivée sous la forme:

$$\begin{aligned} \frac{dL}{dt} &= \sum_i \frac{\partial L}{\partial \dot{q}_i}\ddot{q}_i + \sum_i \frac{d}{dt}\frac{\partial L}{\partial \dot{q}_i}\dot{q}_i \\ &= \frac{d}{dt}\left[\sum_i \frac{\partial L}{\partial \dot{q}_i}\dot{q}_i\right] \ . \end{aligned} \tag{1.66}$$

On trouve ainsi que la quantité

$$\mathcal{E} = \sum_i \frac{\partial L}{\partial \dot{q}_i}\dot{q}_i - L \tag{1.67}$$

est une constante ou intégrale première du mouvement. Notons tout de suite qu'elle a la même dimension que L, c'est à dire celle d'une énergie. Pour préciser son interprétation physique, considérons le cas où les forces internes au système dérivent d'une énergie potentielle U ne dépendant que des q_i. On a alors $L = T(q_i, \dot{q}_i) - U$. T est très généralement une forme quadratique des dérivées temporelles des coordonnées généralisées[8]. T vérifie donc le théorème d'Euler:

$$2T = \sum_i \dot{q}_i \frac{\partial T}{\partial \dot{q}_i} \ . \tag{1.68}$$

Comme U ne dépend que des q_i, $\partial L/\partial \dot{q}_i = \partial T/\partial \dot{q}_i$, et donc

$$\mathcal{E} = 2T - L = 2T - (T - U) = T + U \ . \tag{1.69}$$

La quantité $\mathcal{E}$ coïncide donc dans ce cas simple avec l'énergie mécanique totale du système. Dans un cas plus complexe, nous admettrons donc que l'énergie mécanique du système est définie comme $\mathcal{E}$ et est donc une intégrale première du mouvement. Cette loi de conservation apparaît ici très clairement comme une conséquence directe de l'invariance par translation dans le temps.

L'hypothèse du système isolé exclut bien sûr de traiter le cas d'un ensemble de particules en interaction avec un champ électromagnétique extérieur. Il existe cependant un cas où on peut définir une énergie mécanique conservée pour un tel système: celui d'un champ statique. Les équations de Lagrange s'écrivent en effet normalement avec le lagrangien

$$L = \sum_\alpha T_\alpha - q_\alpha V_\alpha + q_\alpha \mathbf{A}_\alpha \cdot \mathbf{v}_\alpha \ , \tag{1.70}$$

et les notations du paragraphe 1.3 (pour simplifier les écritures nous supposerons ici que les coordonnées généralisées sont les coordonnées cartésiennes). $\mathbf{A}$ ne dépendant que de la position des particules et L

[8]Seulement dans le cas où les contraintes ne font pas explicitement intervenir de dépendance temporelle.

étant explicitement indépendant du temps, le raisonnement précédent s'applique en entier. L'énergie mécanique s'écrit simplement dans ce cas:

$$\mathcal{E} = \sum_\alpha \boldsymbol{\nabla}_{\mathbf{v}_\alpha} L \cdot \mathbf{v}_\alpha - L \ . \tag{1.71}$$

Elle est conservée et on a:

$$\sum_\alpha \boldsymbol{\nabla}_{\mathbf{v}_\alpha} L \cdot \mathbf{v}_\alpha = 2T + \sum_\alpha q_\alpha \boldsymbol{\nabla}_{\mathbf{v}_\alpha} (\mathbf{A}_\alpha \cdot \mathbf{v}_\alpha) \cdot \mathbf{v}_\alpha = 2T + \sum_\alpha q_\alpha \mathbf{A}_\alpha \cdot \mathbf{v}_\alpha \ . \tag{1.72}$$

Et donc:

$$\mathcal{E} = T + \sum_\alpha q_\alpha V_\alpha \ , \tag{1.73}$$

résultat évident de statique.

1.5.2 Translation spatiale: conservation de l'impulsion

Nous considérons, comme dans le paragraphe précédent, un système isolé régi par des équations de Lagrange sous la forme standard. La dynamique est visiblement invariante si on déplace globalement le système. Il n'existe rien dans l'environnement pour fixer une origine des coordonnées. Comme ce déplacement est une quantité vectorielle, on a en fait trois conditions d'invariance d'où on peut déduire la conservation de trois quantités.

Nous allons d'abord supposer que le système est décrit par ses coordonnées cartésiennes. Les coordonnées généralisées posent en effet problème, puisqu'une translation spatiale n'a pas forcément une expression simple en termes de ces coordonnées. Nous examinerons ce cas à la fin du paragraphe. Nous considérons une translation d'ensemble du système, qui s'écrit:

$$\mathbf{r}_\alpha \longrightarrow \mathbf{r}_\alpha + \boldsymbol{\varepsilon} \ . \tag{1.74}$$

La translation $\boldsymbol{\varepsilon}$ est supposée petite à l'échelle des longueurs caractéristiques du système. On peut aisément exprimer la variation de la fonction de Lagrange dans cette translation. Puisque les vitesses ne changent pas,

$$\delta L = \sum_\alpha \boldsymbol{\nabla}_{\mathbf{r}_\alpha} L \cdot \boldsymbol{\varepsilon} \ . \tag{1.75}$$

Cette variation ne peut être nulle quel que soit $\boldsymbol{\varepsilon}$ que si

$$\sum_\alpha \boldsymbol{\nabla}_{\mathbf{r}_\alpha} L = 0 \ . \tag{1.76}$$

En utilisant les équations de Lagrange en coordonnées cartésiennes, on met cette condition sous la forme:

$$\frac{d\mathbf{P}}{dt} = 0 \ , \tag{1.77}$$

avec

$$\mathbf{P} = \sum_\alpha \boldsymbol{\nabla}_{\mathbf{v}_\alpha} L \ . \tag{1.78}$$

Nous trouvons donc bien une intégrale première vectorielle pour le mouvement. Dans le cas où $L = T - U$, U ne dépendant que des $\mathbf{r}_\alpha$, $\boldsymbol{\nabla}_{\mathbf{v}_\alpha} L = \boldsymbol{\nabla}_{\mathbf{v}_\alpha} T$ et

$$\mathbf{P} = \sum_\alpha m_\alpha \mathbf{v}_\alpha \ , \tag{1.79}$$

indiquant que cette intégrale première vectorielle n'est autre que l'impulsion ou quantité de mouvement totale.

Notons à ce point que l'impulsion totale est définie en terme du gradient par rapport aux vitesses et ne coïncide pas forcément avec l'expression habituelle $\sum_\alpha m_\alpha \mathbf{v}_\alpha$. Ce sera en particulier le cas en électromagnétisme où le potentiel vecteur entre dans la définition de l'impulsion. Comme seuls des potentiels vecteur et scalaire uniformes (c'est à dire des champs nuls) seraient compatibles avec les hypothèses actuelles, nous n'examinerons pas ce cas pour l'instant.

Etudions maintenant le cas où l'existence de contraintes impose le recours à des coordonnées généralisées. L'invariance de l'impulsion totale ne devrait pas dépendre de l'existence de contraintes internes au système. Toutefois, la conservation de $\mathbf{P}$ ne peut plus être établie aussi simplement, puisqu'une translation n'a pas nécessairement une expression simple en termes de coordonnées généralisées. Définissons quand même, par analogie avec le cas précédent, une "impulsion généralisée" p_i associée à la coordonnée généralisée q_i par

$$p_i = \frac{\partial L}{\partial \dot{q}_i} \ . \tag{1.80}$$

Nous verrons que des impulsions généralisées jouent un rôle central dans le formalisme hamiltonien. En attendant, notons que l'énergie mécanique totale peut se réécrire simplement:

$$\mathcal{E} = \sum_i p_i \dot{q}_i - L \ . \tag{1.81}$$

Supposons maintenant que la fonction de Lagrange L ne fasse pas explicitement intervenir la coordonnée q_i:

$$\frac{\partial L}{\partial q_i} = 0 \ . \tag{1.82}$$

Nous dirons alors que q_i est une coordonnée "cyclique". La simple écriture de l'équation de Lagrange pour q_i donne alors:

$$\frac{dp_i}{dt} = 0 \ . \tag{1.83}$$

L'impulsion associée à une coordonnée cyclique est conservée.

Revenons maintenant à l'invariance par translation. Trois des coordonnées généralisées d'un système isolé peuvent être prises égales aux trois coordonnées cartésiennes du centre d'inertie. De manière évidente, par invariance par translation, ces trois coordonnées sont cycliques. Il en résulte que les trois impulsions associées sont des intégrales premières du mouvement, formant un vecteur qui n'est autre que l'impulsion totale.

1.5.3 Invariance par rotation: moment cinétique

Nous appliquerons ici une méthode très semblable à celle du paragraphe précédent au cas où le système et donc sa fonction de Lagrange sont invariants dans une rotation quelconque autour d'un axe défini par un vecteur unitaire $\mathbf{u}$. C'est en particulier le cas pour un système isolé. Nous examinerons ici uniquement le cas où le système est défini par les coordonnées cartésiennes des particules.

Nous considérons une rotation infinitésimale d'un angle $\delta\phi$ autour de l'axe. En posant $\delta\boldsymbol{\phi} = \delta\phi\mathbf{u}$, nous pouvons écrire les variations des positions et vitesses dans cette transformation $\delta\mathbf{r} = \delta\boldsymbol{\phi} \times \mathbf{r}$ et $\delta\mathbf{v} = \delta\boldsymbol{\phi} \times \mathbf{v}$. L'invariance par rotation impose que la variation δL de la fonction de Lagrange soit nulle dans cette transformation. Or

$$\delta L = \sum_\alpha \boldsymbol{\nabla}_{\mathbf{r}_\alpha} L \cdot \delta\mathbf{r}_\alpha + \boldsymbol{\nabla}_{\mathbf{v}_\alpha} L \cdot \delta\mathbf{v}_\alpha \ . \tag{1.84}$$

En utilisant les équations de Lagrange, qui s'écrivent $\boldsymbol{\nabla}_{\mathbf{r}_\alpha} L = \dot{\mathbf{p}}_\alpha$, où $\mathbf{p}_\alpha$ est l'impulsion de la particule α, on peut écrire:

$$\delta L \ = \ \sum_\alpha \left[\dot{\mathbf{p}}_\alpha \cdot (\delta\boldsymbol{\phi} \times \mathbf{r}_\alpha) + \mathbf{p}_\alpha \cdot (\delta\boldsymbol{\phi} \times \mathbf{v}_\alpha)\right]$$

$$
\begin{aligned}
&= \sum_\alpha \left[\delta\boldsymbol{\phi}\cdot(\mathbf{r}_\alpha\times\dot{\mathbf{p}}_\alpha)+\delta\boldsymbol{\phi}\cdot(\mathbf{v}_\alpha\times\mathbf{p}_\alpha)\right] \\
&= \delta\boldsymbol{\phi}\cdot\frac{d\mathbf{L}}{dt}\ ,
\end{aligned}
\tag{1.85}
$$

où nous avons posé

$$
\mathbf{L}=\sum_\alpha \mathbf{r}_\alpha\times\mathbf{p}_\alpha\ . \tag{1.86}
$$

La fonction de Lagrange ne peut être conservée pour une rotation arbitraire que si la composante du vecteur $\mathbf{L}$ sur l'axe de rotation est une intégrale première du mouvement. Nous retrouvons ainsi la conservation du moment cinétique par rapport à un axe.

On peut voir aussi, plus simplement, dans ce cas, que le lagrangien est indépendant d'une coordonnée généralisée décrivant le mouvement de rotation autour de l'axe. Il est donc cyclique dans cette coordonnée. L'impulsion généralisée associée à cette coordonnée et constante. Il est facile de vérifier qu'il s'agit bien de la composante le long de cet axe du moment cinétique.

Dans le cas où le système est invariant dans une rotation arbitraire autour d'un axe quelconque, ce qui est le cas d'un système isolé, il en résulte que le moment cinétique $\mathbf{L}$ est une constante du mouvement.

1.6 Action en fonction de la trajectoire

Nous allons utiliser les définitions du paragraphe précédent pour tenter d'exprimer simplement la dépendance de l'action S calculée sur la trajectoire effectivement suivie en fonction des coordonnées spatiales et temporelles du point de départ et du point d'arrivée. En d'autres termes, l'objet de ce paragraphe est de donner les dérivées partielles de $S(q(1),t_1,q(2),t_2)$, action sur la trajectoire suivie considérée comme une fonction des conditions aux limites. Ces résultats nous seront fort utiles dans la suite du cours. Ils nous permettront aussi de jeter un regard nouveau sur les lois de conservation associées aux invariances.

1.6.1 Dépendance en position

Nous considérons ici la variation de S en fonction de la position du point d'arrivée. Nous en déduirons immédiatement la dépendance vis à vis du point de départ. Nous considérons donc ici deux trajectoires *effectivement suivies* entre les instants t_1 et t_2. La première (trajectoire de référence) s'effectue entre les valeurs $q_i(1)$ et $q_i(2)$ des coordonnées généralisées. La seconde entre les valeurs $q_i(1)$ et $q_i(2)+\delta q_i(2)$, où $\delta q_i(2)$ est un accroissement infinitésimal de la position du point d'arrivée. Les trajectoires, par continuité, restent infiniment voisines pour tous les instants[9] et on écrira $\delta q_i(t)$ l'écart entre elles.

La variation de l'action en passant d'une trajectoire à l'autre, δS, s'écrit simplement:

$$
\delta S=\int_{t_1}^{t_2}\delta L\,dt\ , \tag{1.87}
$$

où δL est la variation de la fonction de Lagrange entre les deux points à l'instant t. On suit alors un raisonnement identique à celui utilisé pour établir les équations de Lagrange. On écrit d'abord:

$$
\delta L=\sum_i\frac{\partial L}{\partial q_i}\delta q_i+\sum_i\frac{\partial L}{\partial\dot{q}_i}\delta\dot{q}_i\ , \tag{1.88}
$$

[9] Si la dynamique du système était chaotique, deux trajectoires très voisines aux points de départ et d'arrivée, peuvent s'écarter notablement l'une de l'autre. Comme nous manipulons des accroissements infinitésimaux, ce problème ne se pose pas.

et on reporte cette expression dans celle de δS. Le terme faisant intervenir $\delta\dot{q}_i$ est alors intégré par parties. On obtient:

$$\delta S = \int_{t_1}^{t_2} \sum_i \left[\frac{\partial L}{\partial q_i} - \frac{d}{dt}\frac{\partial L}{\partial \dot{q}_i}\right] \delta q_i \, dt + \sum_i \left[\frac{\partial L}{\partial \dot{q}_i}\delta q_i\right]_1^2 . \tag{1.89}$$

La trajectoire de référence que nous considérons ici est une solution des équations du mouvement. Les équations de Lagrange étant constamment vérifiées, l'intégrale dans l'expression ci-dessus s'annule identiquement. Il ne reste donc que le terme tout intégré, qui se réduit à:

$$\delta S = \sum_i \frac{\partial L}{\partial \dot{q}_i}\delta q_i(2) . \tag{1.90}$$

De manière évidente, si nous avions considéré deux trajectoires différant d'une quantité infinitésimale au point de départ, nous aurions obtenu:

$$\delta S = -\sum_i \frac{\partial L}{\partial \dot{q}_i}\delta q_i(1) . \tag{1.91}$$

Ces deux expressions nous donnent les dérivées partielles de l'action, considérée comme une fonction des coordonnées des points de départ et d'arrivée. En remarquant que $\partial L/\partial \dot{q}_i = p_i$ (impulsion généralisée), on écrira

$$\frac{\partial S}{\partial q_i(2)} = p_i(2) \tag{1.92}$$

$$\frac{\partial S}{\partial q_i(1)} = -p_i(1) \tag{1.93}$$

On peut facilement retrouver, à partir de ce résultat, l'invariance de l'impulsion associée à une coordonnée cyclique. Si q_i est cyclique, la fonction de Lagrange ne dépend pas de q_i et le mouvement doit être invariant dans une translation de la coordonnée q_i. Considérons donc la translation infinitésimale $q_i \to q_i + \epsilon$. La variation de l'action dans cette transformation est:

$$\delta S = \frac{\partial S}{\partial q_i(2)}\epsilon + \frac{\partial S}{\partial q_i(1)}\epsilon = (p_i(2) - p_i(1))\epsilon , \tag{1.94}$$

d'où on tire immédiatement la conservation de p_i. Dans le cas où les coordonnées cycliques sont celles du centre d'inertie, on retrouve la conservation de l'impulsion au sens habituel.

1.6.2 Dépendance en temps

Nous considérons maintenant la dépendance de l'action dans le temps d'arrivée t_2. Nous considérons donc deux trajectoires effectivement suivies par le système. L'une, entre $q(1), t_1$ et $q(2), t_2$ est la trajectoire de référence. L'autre coïncide avec la première jusqu'à l'instant t_2 et continue ensuite pendant un intervalle de temps infinitésimal jusqu'à $t_2 + \delta t_2$, les coordonnées étant alors $q_i(2) + \delta q_i(2)$. On peut écrire de manière évidente la variation de l'action entre ces deux trajectoires comme:

$$\delta S = L(t_2)\delta t_2 , \tag{1.95}$$

mais aussi comme:

$$\delta S = \frac{\partial S}{\partial t_2}\delta t_2 + \sum_i \frac{\partial S}{\partial q_i(2)}\delta q_i(2) . \tag{1.96}$$

En effet l'instant du point d'arrivée et ses coordonnées varient. En utilisant les résultats du paragraphe précédent et en remarquant que $\delta q_i(2) = \dot{q}_i(2)\delta t_2$, on peut écrire:

$$L(t_2) = \frac{\partial S}{\partial t_2} + \sum_i p_i(2)\dot{q}_i(2) . \tag{1.97}$$

En nous souvenant de la définition de l'énergie mécanique totale $\mathcal{E}$[10], nous avons enfin:

$$\frac{\partial S}{\partial t_2} = -\mathcal{E}(2) \ . \tag{1.98}$$

En reproduisant le même raisonnement pour une variation de l'instant de départ, on obtiendrait de même:

$$\frac{\partial S}{\partial t_1} = \mathcal{E}(1) \ . \tag{1.99}$$

Avec ces deux expressions et les résultats du paragraphe précédent, nous connaissons toutes les dérivées partielles de l'action en fonction des conditions aux limites imposées à la trajectoire.

Notons également qu'on peut retrouver facilement à partir de ces expressions la conservation de l'énergie mécanique. Si la fonction de Lagrange ne dépend pas explicitement du temps, l'action doit être invariante dans une translation temporelle globale infinitésimale. La variation de l'action dans cette translation s'écrivant évidemment $\delta S = (\mathcal{E}(1) - \mathcal{E}(2))\delta t$, on retrouve l'invariance de l'énergie mécanique.

1.7 Deux théorèmes utiles

Nous considérerons dans ce paragraphe le cas particulier important des énergies potentielles homogènes, c'est à dire telles que:

$$U(\alpha q_1, \ldots, \alpha q_n) = \alpha^k U(q_1, \ldots, q_n) \ . \tag{1.100}$$

Ce cas, très spécifique, se rencontre en fait dans une grande variété de problèmes de mécanique:

- Pour l'oscillateur harmonique unidimensionnel $U = Kx^2/2$ est évidemment une fonction homogène avec $k = 2$. Ce résultat est bien sûr valide aussi pour les oscillateurs multidimensionnels.
- Le potentiel gravitationnel ou Coulombien en $1/r$ correspond à $k = -1$.
- Le mouvement dans un champ uniforme (par exemple champ de pesanteur local) correspond enfin à $k = 1$.

Le fait que U soit une fonction homogène a deux conséquences importantes que nous allons examiner dans les prochains paragraphes. D'abord, on peut en déduire des lois d'échelle utiles fixant la dépendance relative de certains paramètres des trajectoires (par exemple la troisième loi de Kepler). La deuxième conséquence est le théorème du viriel, d'une grande importance en mécanique et physique statistique.

1.7.1 Lois d'échelle

Considérons une transformation d'échelle sur les coordonnées généralisées et le temps. Elle peut s'écrire:

$$q_i \longrightarrow \alpha q_i \tag{1.101}$$

$$t \longrightarrow \beta t \tag{1.102}$$

[10] Nous n'avons fait dans ce paragraphe aucune hypothèse sur la dépendance en temps de la fonction de Lagrange. $\mathcal{E}$ n'est donc pas nécessairement une quantité conservée. En fait, plutôt que d'énergie totale, nous devrions parler de fonction de Hamilton (voir chapitre suivant).

On en déduit les transformations de quelques quantités:

$$\dot{q}_i \longrightarrow \frac{\alpha}{\beta}\dot{q}_i \tag{1.103}$$

$$U \longrightarrow \alpha^k U \tag{1.104}$$

$$T \longrightarrow \left(\frac{\alpha}{\beta}\right)^2 T\,. \tag{1.105}$$

Pour établir la dernière ligne, il faut admettre que l'énergie cinétique ne dépend que des $\dot{q}_i$ et pas de q_i. C'est en particulier le cas en coordonnées cartésiennes standard. Nous supposerons cette condition remplie dans toute la suite du paragraphe.

En général, le lagrangien L ne se transforme pas de manière simple. En revanche, dans le cas particulier où les facteurs d'échelle sur T et U sont les mêmes, c'est à dire si on choisit les facteurs d'échelle α et β de telle manière que:

$$\left(\frac{\alpha}{\beta}\right)^2 = \alpha^k\,, \tag{1.106}$$

où encore

$$\beta = \alpha^{1-k/2}\,, \tag{1.107}$$

la fonction de Lagrange est simplement multipliée par un facteur d'échelle. Un tel facteur est sans effet sur les équations du mouvement. A une trajectoire possible correspond donc après cette transformation d'espace et de temps une autre trajectoire possible. Cette correspondance permet par exemple d'obtenir les lois reliant la période d'un mouvement à l'extension spatiale de la trajectoire. Appliquons ces arguments aux cas particulièrement importants de l'oscillateur harmonique et du mouvement gravitationnel.

- Pour l'oscillateur harmonique, $k = 2$. L'équation (1.107) est donc satisfaite pour n'importe quelle valeur de α à condition que $\beta = 1$. On obtient donc une trajectoire possible en multipliant la coordonnée spatiale par un facteur arbitraire et en laissant inchangée l'échelle de temps. On retrouve ainsi très simplement l'isochronisme des oscillations de l'oscillateur harmonique.

- Pour le mouvement dans un champ gravitationnel uniforme, on a $k = 1$. L'équation (1.107) est donc satisfaite si $\beta = \sqrt{\alpha}$. Deux trajectoires se correspondent donc si leurs extensions spatiales ℓ et ℓ' et temporelles ($\mathcal{T}$ et $\mathcal{T}'$) sont reliées par $\mathcal{T}'/\mathcal{T} = \sqrt{\ell'/\ell}$. On retrouve ainsi, par exemple, la dépendance en $\sqrt{\ell}$ de la période des oscillations d'un pendule simple de longueur ℓ.

- Enfin, pour le cas du mouvement de Kepler dans un potentiel gravitationnel central, on a $k = -1$. La relation (1.107) est donc satisfaite si $\beta = \alpha^{3/2}$. On en déduit alors, par exemple, la troisième loi de Kepler qui relie la période d'une orbite $\mathcal{T}$ à son demi grand axe a: $\mathcal{T}^2/a^3$ est une constante.

Ces trois exemples illustrent bien la puissance de ces simples lois d'échelle. En fait, ce genre d'arguments se transpose à des domaines très variés et permet souvent d'obtenir des lois très générales par de simples considérations d'échelle ou de dimensionnalité.

1.7.2 Théorème du Viriel

Nous établirons ici, dans le cas des énergies potentielles homogènes, un lien utile entre les énergies potentielles et cinétiques moyennes (moyennées sur un temps long devant les temps caractéristiques du mouvement, la période par exemple). Ce théorème possède de nombreuses applications en mécanique céleste.

L'énergie cinétique T est une forme quadratique des vitesses. On peut donc lui appliquer le théorème d'Euler:

$$2T = \sum_i \frac{\partial T}{\partial \dot{q}_i}\dot{q}_i\ . \tag{1.108}$$

L'énergie potentielle ne dépendant pas des vitesses généralisées, $\partial T/\partial \dot{q}_i = \partial L/\partial \dot{q}_i$. En utilisant la définition des impulsions généralisées, on peut mettre cette dernière relation sous la forme:

$$2T = \sum_i p_i \dot{q}_i = \frac{d}{dt} \sum_i p_i q_i - \sum_i \dot{p}_i q_i \ . \tag{1.109}$$

Prenons maintenant la valeur moyenne temporelle de T définie comme:

$$\overline{T} = \lim_{\tau \to \infty} \frac{1}{\tau} \int_0^\tau T \, dt \ . \tag{1.110}$$

Notons que, dans le cas des mouvements périodiques, on peut définir la moyenne temporelle comme l'intégrale sur une période. En incluant dans cette définition de la valeur moyenne l'expression précédente de T, on trouve:

$$2\overline{T} = \lim_{\tau \to \infty} \frac{1}{\tau} \left[\sum_i p_i q_i \right]_0^\tau - \frac{1}{\tau} \int_0^\tau \sum_i \dot{p}_i q_i \ . \tag{1.111}$$

Si le mouvement est *borné*, le terme tout intégré reste fini quand τ tend vers l'infini et ne contribue donc pas à la valeur moyenne (il est strictement nul si le mouvement est périodique et si les intégrales s'étendent sur exactement une période). On a donc:

$$2\overline{T} = -\sum_i \overline{\dot{p}_i q_i} \ , \tag{1.112}$$

ou encore, en utilisant les équations de Lagrange $\dot{p}_i = \partial L/\partial q_i$,

$$2\overline{T} = -\sum_i \overline{\frac{\partial L}{\partial q_i} q_i} \ . \tag{1.113}$$

Si, comme dans le paragraphe précédent, T ne dépend que des $\dot{q}_i$, on a $\partial L/\partial q_i = -\partial U/\partial q_i$. Si enfin U est une fonction homogène des coordonnées de degré k, on peut écrire:

$$2\overline{T} = k\overline{U} \ . \tag{1.114}$$

Rappelons pour finir les hypothèses nécessaires à l'établissement de cette propriété:

- Les forces dérivent d'une énergie potentielle et les équations de Lagrange s'appliquent sous la forme ordinaire.
- L'énergie potentielle est une fonction des coordonnées homogène et de degré k.
- L'énergie cinétique ne dépend que des vitesses généralisées.
- Le mouvement est borné.

Ces hypothèses contraignantes sont en fait assez souvent remplies. C'est par exemple le cas pour les mouvements périodiques dans un potentiel gravitationnel central, où la relation s'écrit, avec $k = -1$:

$$\overline{T} = -\frac{1}{2}\overline{U} \tag{1.115}$$

ou pour les mouvements dans un potentiel harmonique (nécessairement bornés) pour lesquels $k = 2$ et

$$\overline{T} = \overline{U} \ . \tag{1.116}$$

Chapitre 2

Formulation hamiltonienne

Nous présenterons brièvement dans ce chapitre la formulation hamiltonienne de la mécanique. Bien qu'elle soit strictement équivalente à la formulation lagrangienne, elle s'avère souvent plus puissante et d'un usage plus commode.

D'abord, d'un point de vue très utilitaire, la formulation hamiltonienne facilite la résolution, analytique ou numérique, des équations du mouvement. Les équations de Lagrange sont des équations différentielles du second ordre. Leur résolution analytique n'est pas toujours facile, si elle est possible. Très souvent (par exemple dès le problème à trois corps en mécanique céleste), on doit recourir à une intégration numérique. Il se trouve que les équations différentielles du second ordre ne se prêtent pas très bien à une telle intégration. Il se pose en particulier des problèmes sévères de stabilité numérique. En revanche, les équation différentielles du premier ordre s'intègrent très aisément et de façon numériquement stable. Comme nous le verrons, les équations de Hamilton sont du premier ordre.

En fait, le formalisme hamiltonien consiste à traiter sur un même plan les positions et les impulsions généralisées associées. Outre la transformation d'une équation de Lagrange du second ordre, portant sur la position seule, en deux équations de Hamilton du premier ordre reliant position et impulsion, cette approche ouvre la voie à des changements de variables (nous dirons des transformations canoniques) très puissants. Plutôt que de changer simplement de coordonnées généralisées, on peut en effet mélanger positions et impulsions dans un changement de variables. Il est même possible ainsi d'échanger le rôle des impulsions et des positions ou de prendre les conditions aux limites comme nouvelles variables. Cette immense liberté dans le choix de la description du problème est au centre de nombreuses méthodes de résolution des problèmes de mécanique.

Un autre avantage décisif de la formulation hamiltonienne est qu'elle se prête à merveille à des méthodes de perturbations (développement de la solution en puissance des perturbations, supposées petites, par rapport à un problème dont la solution est connue). Ces méthodes sont d'une importance primordiale en mécanique céleste (prise en compte, par exemple, des perturbations au mouvement de Kepler dues aux autres planètes). Enfin, et surtout, c'est la formulation hamiltonienne de la mécanique classique qui se prête à la quantification (dite, elle aussi, canonique). Nous n'aborderons pas ce problème, mais un certain nombre des notions que nous introduirons dans ce chapitre parfaitement classique ont une contrepartie dans le formalisme quantique. Si, de manière évidente, la fonction de Hamilton est remplacée par l'opérateur hamiltonien, les crochets de Poisson, par exemple, correspondent aux commutateurs.

2.1 Equations de Hamilton

Nous partons de la formulation lagrangienne établie au chapitre précédent, dans sa forme la plus simple. Nous n'intégrerons donc pas les généralisations aux forces ne dérivant pas d'un potentiel (mises à part les forces électromagnétiques) ni les généralisations à des liaisons non holonomes. On

définit les impulsions généralisées à partir de la fonction de Lagrange des positions, de vitesses et du temps[1] par

$$p_i = \frac{\partial L}{\partial \dot{q}_i} \tag{2.1}$$

et on peut écrire les équations de Lagrange sous la forme:

$$\dot{p}_i = \frac{\partial L}{\partial q_i} \ . \tag{2.2}$$

Ces écritures font apparaître L comme une fonction naturelle des q_i et des $\dot{q}_i$, dont les dérivées partielles respectives sont les $\dot{p}_i$ et les p_i. La "différentielle totale" du lagrangien s'écrit alors:

$$dL = \sum_i p_i d\dot{q}_i + \sum_i \dot{p}_i dq_i \ . \tag{2.3}$$

La technique des transformations de Legendre, largement illustrée en thermodynamique classique[2], permet de passer de L à une fonction dont la différentielle s'exprime de façon naturelle en fonction des q_i et des p_i. Posons en effet

$$H = \sum_i p_i \dot{q}_i - L \ . \tag{2.4}$$

Nous appellerons H la fonction de Hamilton. Notons que, pour les systèmes conservatifs, la définition de la fonction de Hamilton coïncide avec celle de l'énergie totale. En revanche, pour des systèmes où la fonction de Lagrange dépend explicitement du temps, la notion d'énergie mécanique totale perd tout son intérêt, alors que la fonction de Hamilton garde son sens et permet toujours d'écrire les équations du mouvement. C'est ce caractère plus général de la fonction de Hamilton qui justifie de ne pas la confondre avec l'énergie totale. La différentielle de la fonction de Hamilton s'écrit sans difficulté comme:

$$dH = \sum_i \dot{q}_i dp_i - \sum_i \dot{p}_i dq_i \ . \tag{2.5}$$

H apparaît donc comme une fonction naturelle des p_i et q_i dont les dérivées partielles respectives sont les $\dot{q}_i$ et les $\dot{p}_i$. Les $2n$ équations de Hamilton, équivalentes aux n équations de Lagrange, qui permettent la résolution du problème avec les conditions aux limites, s'écrivent donc:

$$\frac{\partial H}{\partial p_i} = \dot{q}_i \tag{2.6}$$

$$\frac{\partial H}{\partial q_i} = -\dot{p}_i \ , \tag{2.7}$$

auxquelles on pourrait ajouter la relation $\partial H/\partial t = -\partial L/\partial t$, qui ne présente d'intérêt que pour un lagrangien et donc un hamiltonien dépendant explicitement du temps. Les positions et les impulsions sont dites variables conjuguées.

Comme attendu, nous avons donc remplacé les n équations différentielles du second ordre que nous donne le formalisme lagrangien, par un nombre double d'équations différentielles du premier ordre qui se prêtent plus facilement à une résolution analytique ou numérique. Les conditions initiales naturelles pour ces équations sont les positions et les impulsions initiales (alors que c'étaient les positions et les vitesses initiales pour les équations de Lagrange). Rappelons que les impulsions ne coïncident pas nécessairement avec les vitesses généralisées, en particulier en présence de champs électromagnétiques).

[1] Nous supposerons très souvent dans ce chapitre que la fonction de Lagrange ne dépend pas explicitement du temps. L'énergie mécanique totale est alors une intégrale première du mouvement.

[2] C'est en effet la transformation qui permet, par exemple, de passer de l'énergie interne U, dont la différentielle s'exprime simplement en fonction de celles du volume et de l'entropie, à l'enthalpie H, faisant intervenir pression et entropie.

Notons enfin la symétrie remarquable de ces équations, qui traitent de façon tout à fait analogue les positions et les impulsions (à un signe près).

Notons aussi que si q_i est une coordonnée cyclique, la fonction de Lagrange et donc la fonction de Hamilton ne dépendent pas explicitement de q_i. On déduit alors immédiatement des équations de Hamilton que p_i est une constante du mouvement.

Examinons maintenant le cas particulier où les forces dérivent d'une énergie potentielle ne dépendant que des positions et où les coordonnées généralisées coïncident avec les coordonnées cartésiennes. La fonction de Lagrange s'écrit $L = T - U$ et les impulsions généralisées coïncident avec la définition habituelle $\mathbf{p}_\alpha = m_\alpha \mathbf{v}_\alpha$. La fonction de Hamilton, qui est identique à l'énergie totale, s'écrit alors $H = T + U$ et les équations de Hamilton s'écrivent:

$$\boldsymbol{\nabla}_{\mathbf{p}_\alpha} H = \mathbf{v}_\alpha \tag{2.8}$$

$$\boldsymbol{\nabla}_{\mathbf{r}_\alpha} H = -\dot{\mathbf{p}}_\alpha \ . \tag{2.9}$$

En écrivant $H = \sum_\alpha \mathbf{p}_\alpha^2/2m_\alpha + U$, on en déduit:

$$\mathbf{p}_\alpha/m_\alpha = \mathbf{v}_\alpha \tag{2.10}$$

$$\mathbf{F}_\alpha = \dot{\mathbf{p}}_\alpha \ . \tag{2.11}$$

La première de ces deux équations de Hamilton n'est autre que la définition de l'impulsion en termes de la vitesse. La deuxième n'est autre que le principe fondamental de la dynamique, où on a remplacé l'accélération par la dérivée par rapport au temps de l'impulsion. Dans ce cas très simple, l'écriture des équations de Hamilton à partir des équations de Newton revient à prendre comme variables indépendantes position et vitesse, un changement de variable bien connu pour transformer les équations différentielles du second ordre en équations du premier ordre.

Examinons maintenant un autre cas très important: celui de particules chargées dans un champ électromagnétique. Nous supposerons encore que les coordonnées sont les coordonnées cartésiennes habituelles. Nous avons montré au chapitre précédent que la fonction de Lagrange s'écrit:

$$L = \sum_\alpha \left(T_\alpha - q_\alpha V_\alpha + q_\alpha \mathbf{A}_\alpha \cdot \mathbf{v}_\alpha \right) \ , \tag{2.12}$$

où V_α et $\mathbf{A}_\alpha$ sont les potentiels scalaires et vecteur vus par la particule α.

Pour passer aux équations de Hamilton, la première étape est de déterminer l'impulsion généralisée $\mathbf{p}_\alpha$:

$$\mathbf{p}_\alpha = \boldsymbol{\nabla}_{\mathbf{v}_\alpha} L = m_\alpha \mathbf{v}_\alpha + q_\alpha \mathbf{A}_\alpha \ . \tag{2.13}$$

On voit bien que, sauf dans le cas de l'électrostatique, l'impulsion généralisée ne coïncide pas avec la quantité de mouvement ordinaire.

On peut alors écrire sans difficultés la fonction de Hamilton:

$$H = \sum_\alpha \mathbf{p}_\alpha \cdot \mathbf{v}_\alpha - L = \sum_\alpha \left(\frac{1}{2} m_\alpha \mathbf{v}_\alpha^2 + q_\alpha V_\alpha \right) \ . \tag{2.14}$$

De façon très remarquable, la fonction de Hamilton écrite sous cette forme ne fait pas intervenir le potentiel vecteur magnétique. En effet, au moins dans le cas où les potentiels sont statiques, nous avons vu au chapitre précédent que l'énergie totale conservée des particules est la somme de leur énergie cinétique et de l'énergie potentielle électrostatique. On pourrait craindre qu'une telle fonction de Hamilton ne conduise à des équations du mouvement qui ne font plus intervenir le champ magnétique. Bien sûr il n'en est rien. On doit en effet, pour écrire les équations de Hamilton, exprimer H comme une fonction des impulsions généralisées, non pas des vitesses. En utilisant le lien entre impulsion et vitesse, on obtient:

$$H = \sum_\alpha \left[\frac{(\mathbf{p}_\alpha - q_\alpha \mathbf{A}_\alpha)^2}{2m_\alpha} + q_\alpha V_\alpha \right] \ . \tag{2.15}$$

Cette expression du hamiltonien de particules dans un champ nous sera utile dans la quatrième partie. Nous aurons en effet besoin du hamiltonien quantique d'un atome dans un champ de rayonnement. Nous utiliserons simplement l'expression précédente, en remplaçant toutes les quantités relatives à la particule par les opérateurs quantiques correspondants.

A titre d'exercice, écrivons les équations du mouvement à partir de cette expression du hamiltonien, et vérifions que nous retrouvons bien la force de Lorentz habituelle. Pour alléger les notations, nous ne considérerons que le cas d'une particule unique et nous omettrons donc l'indice α. Le lecteur pourra aisément rétablir le cas général.

Notons d'abord qu'une des équations de Hamilton redonne trivialement, comme dans le cas d'un potentiel standard, la définition de l'impulsion généralisée en termes de la vitesse. Nous nous focaliserons donc sur l'autre équation, qui s'écrit:

$$\dot{\mathbf{p}} = -\boldsymbol{\nabla} H = -q\boldsymbol{\nabla} V - \boldsymbol{\nabla}\frac{(\mathbf{p}-q\mathbf{A})^2}{2m} \ , \tag{2.16}$$

où les gradients s'entendent par rapport à la position $\mathbf{r}$ de la particule. Le gradient du carré scalaire peut se développer comme gradient d'un produit scalaire:

$$\boldsymbol{\nabla}(\mathbf{p}-q\mathbf{A})^2 = 2\left\{\left[(\mathbf{p}-q\mathbf{A})\cdot\boldsymbol{\nabla}\right](\mathbf{p}-q\mathbf{A}) + (\mathbf{p}-q\mathbf{A})\times\left[\boldsymbol{\nabla}\times(\mathbf{p}-q\mathbf{A})\right]\right\} \ . \tag{2.17}$$

Dans cette équation, seul le potentiel vecteur est à considérer comme une fonction de $\mathbf{r}$. On peut alors simplifier cette expression et la dérivée de l'impulsion s'écrit, en faisant intervenir la vitesse de la particule:

$$\dot{\mathbf{p}} = -\boldsymbol{\nabla} H = -q\boldsymbol{\nabla} V + q(\mathbf{v}\cdot\boldsymbol{\nabla})\mathbf{A} + q\mathbf{v}\times\mathbf{B} \ . \tag{2.18}$$

En nous souvenant que:

$$-\boldsymbol{\nabla} V = \mathbf{E} + \frac{\partial \mathbf{A}}{\partial t} \tag{2.19}$$

et que

$$(\mathbf{v}\cdot\boldsymbol{\nabla})\mathbf{A} = \frac{d\mathbf{A}}{dt} - \frac{\partial \mathbf{A}}{\partial t} \tag{2.20}$$

(dérivée hydrodynamique), on a

$$\dot{\mathbf{p}} = q\mathbf{E} + q\mathbf{v}\times\mathbf{B} + q\frac{d\mathbf{A}}{dt} \ . \tag{2.21}$$

Si enfin on exprime la variation de l'impulsion en fonction de l'accélération, on retrouve bien le principe fondamental avec la force de Lorentz sous sa forme standard. Si nous ne doutions guère du résultat, il était important de l'établir explicitement. Il nous a fait bien sentir la différence essentielle entre quantité de mouvement (au sens de la dynamique newtonienne) et impulsion généralisée.

2.2 Crochets de Poisson

Nous allons introduire dans ce paragraphe une notation que nous n'utiliserons guère dans ce cours, très élémentaire, de mécanique analytique. En revanche, l'analogie et le lien formel très important entre ces crochets de Poisson et les commutateurs de la mécanique quantique rendent importante leur introduction à ce point. Nous préciserons d'ailleurs rapidement ces liens.

Considérons une fonction f quelconque des impulsions, des positions et du temps. On peut en écrire la dérivée totale par rapport au temps, le long de la trajectoire suivie par le système, sous la forme:

$$\frac{df}{dt} = \frac{\partial f}{\partial t} + \sum_i \frac{\partial f}{\partial p_i}\dot{p}_i + \sum_i \frac{\partial f}{\partial q_i}\dot{q}_i \ . \tag{2.22}$$

En utilisant les équations de Hamilton, nous mettrons cette dérivée sous la forme:

$$\frac{df}{dt} = \frac{\partial f}{\partial t} + \{H, f\} \ , \tag{2.23}$$

où $\{H, f\}$ est le crochet de Poisson de f avec H, défini par:

$$\{H, f\} = \sum_i \frac{\partial H}{\partial p_i}\frac{\partial f}{\partial q_i} - \frac{\partial H}{\partial q_i}\frac{\partial f}{\partial p_i} \ . \tag{2.24}$$

Le crochet de Poisson est une forme bilinéaire antisymétrique de ses arguments. C'est aussi le cas des commutateurs entre opérateurs en mécanique quantique. Au delà d'une simple remarque mathématique et d'une évidente analogie de notations, nous allons, tout au long de ce paragraphe, rencontrer de nombreuses similitudes entre crochets de Poisson et commutateurs. En fait les uns sont la version classique des autres.

Notons également que les crochets de Poisson peuvent être utilisés pour décrire l'évolution temporelle d'une densité de probabilité dans l'espace des phases. En effet, si $W(p_i, q_i)$ représente la probabilité pour que les coordonnées généralisées et leurs impulsions conjuguées prennent les valeurs p_i, q_i, W évolue dans le temps selon l'équation:

$$\frac{\partial W}{\partial t} = \{H, W\} \ . \tag{2.25}$$

Cette équation est en fait le premier terme d'un développement en puissances de $\hbar$ de l'évolution d'une fonction de distribution de quasi–probabilité dans l'espace des phase, très utilisée en mécanique quantique: la distribution de Wigner (développement de Moyal). Au premier ordre en $\hbar$, comme on pouvait s'y attendre, l'évolution quantique de la densité de probabilité dans l'espace des phases est la même que l'équation d'évolution classique.

Supposons que la fonction f ne fasse pas intervenir explicitement le temps. La condition nécessaire et suffisante pour que f soit une intégrale première du mouvement est alors que $\{H, f\} = 0$. La nullité du crochet de Poisson avec la fonction de Hamilton est équivalente avec la constance de la fonction. Cette propriété est à rapprocher de son équivalent quantique. La condition pour qu'un opérateur F soit une constante du mouvement est en effet que son commutateur $[H, F]$ avec le hamiltonien s'annule.

On peut bien sûr définir le crochet de Poisson de deux fonctions quelconques des positions des impulsions et du temps, estimées sur la trajectoire du système:

$$\{f, g\} = \sum_i \frac{\partial f}{\partial p_i}\frac{\partial g}{\partial q_i} - \frac{\partial f}{\partial q_i}\frac{\partial g}{\partial p_i} \ . \tag{2.26}$$

On déduit de cette expression les propriétés essentielles de l'algèbre des crochets de Poisson:

$$\{f, g\} = -\{g, f\} \tag{2.27}$$
$$\{f, C\} = 0 \tag{2.28}$$
$$\{f + f', g\} = \{f, g\} + \{f', g\} \tag{2.29}$$
$$\{Cf, g\} = C\{f, g\} \tag{2.30}$$
$$\{f, q_i\} = \frac{\partial f}{\partial p_i} \tag{2.31}$$
$$\{f, p_i\} = -\frac{\partial f}{\partial q_i} \ , \tag{2.32}$$

où C est une constante arbitraire. On en déduit facilement les crochets de Poisson des impulsions et positions:

$$\{q_i, q_k\} = 0 \tag{2.33}$$
$$\{p_i, p_k\} = 0 \tag{2.34}$$
$$\{p_i, q_k\} = \delta_{i,k} \tag{2.35}$$

Il est intéressant de noter la similitude entre ces "relations de commutation" classiques et leur analogue quantique ($[X_i, P_k] = i\hbar\delta_{ik}$). De façon très générale, nous verrons que les crochets de Poisson des quantités classiques coïncident, à un facteur $-i\hbar$ près, avec les commutateurs quantiques.

Notons une dernière propriété utile des crochets de Poisson, la relation de Jacobi:

$$\{f, \{g, k\}\} + \{g, \{k, f\}\} + \{k, \{f, g\}\} = 0 \tag{2.36}$$

(nullité de la somme de tous les crochets de Poisson obtenus par permutation circulaire des trois fonctions).

Comme nous l'avons déjà mentionné, les crochets de Poisson sont utiles pour la recherche des intégrales premières du mouvement. Précisons encore ce point en montrant que si f et g sont deux intégrales premières du mouvement, leur crochet de Poisson est aussi une intégrale première. En d'autres termes, nous allons montrer que, si $df/dt = 0$ et $dg/dt = 0$, alors $d\{f, g\}/dt = 0$. Nous avons:

$$\begin{aligned}
\frac{d}{dt}\{f, g\} &= \frac{\partial}{\partial t}\{f, g\} + \{H, \{f, g\}\} \\
&= \left\{\frac{\partial f}{\partial t}, g\right\} + \left\{f, \frac{\partial g}{\partial t}\right\} - \{f, \{g, H\}\} - \{g, \{H, f\}\} \\
&= \left\{\frac{\partial f}{\partial t} + \{H, f\}, g\right\} + \left\{f, \frac{\partial g}{\partial t} + \{H, g\}\right\} \\
&= \left\{\frac{df}{dt}, g\right\} + \left\{f, \frac{dg}{dt}\right\} \\
&= 0 \ .
\end{aligned} \tag{2.37}$$

Si donc on connaît deux intégrales premières du mouvement, on peut en trouver en principe une troisième en prenant leur crochet de Poisson. Rien ne garantit néanmoins que cette troisième intégrale ne soit triviale (nulle en particulier) ou déjà connue. Notons là encore qu'il existe un analogue quantique évident à cette propriété: si deux opérateurs commutent avec le hamiltonien (et sont donc des constantes du mouvement), alors leur commutateur commute lui aussi avec le hamiltonien et donne une troisième constante du mouvement.

Pour clore ce paragraphe, considérons brièvement le cas des trois composantes L_x, L_y et L_z du moment cinétique total $\mathbf{L}$. Nous nous placerons, pour fixer les idées, dans le cas d'une particule unique, et nous laisserons les généralisations au lecteur. Nous avons montré au chapitre précédent que $\mathbf{L} = \mathbf{r} \times \mathbf{p}$. On en déduit $L_x = yp_z - zp_y$ (et les deux autres composantes par permutation circulaire des indices). Le crochet de Poisson $\{L_x, L_y\}$ fait intervenir les dérivées partielles des composantes concernées du moment cinétique par rapport aux composantes de la position et de l'impulsion. L'expression des composantes du moment cinétique permet de calculer facilement ces dérivées. Après un calcul sans grand intérêt, on obtient:

$$\{L_x, L_y\} = -L_z \tag{2.38}$$

et les trois relations se déduisant de celle ci par permutation circulaire des indices. Notons, une fois de plus, l'analogie entre ces équations et leur contrepartie quantique:

$$[L_x, L_y] = i\hbar L_z \ , \tag{2.39}$$

relations d'un très grande importance puisqu'elles définissent ce qu'est un moment cinétique quantique.

Nous avons vu au chapitre précédent que l'invariance par rotation arbitraire autour d'un axe impliquait la conservation de la composante du moment cinétique sur cet axe. Imaginons qu'on ait pu montrer la conservation de L_x et L_y. Il en résulte, comme nous venons de le montrer, la conservation de leur crochet de Poisson, c'est à dire de L_z. Il suffit en fait que le moment par rapport à deux axes perpendiculaires soit conservé pour que le moment cinétique total le soit.

Nous pouvons aussi appliquer ces relations à l'étude d'un problème dynamique important. Nous considérerons un système dynamique dont le hamiltonien peut s'écrire $H = \mathbf{\Omega} \cdot \mathbf{L}$, où $\mathbf{\Omega}$ est un vecteur constant. C'est par exemple le hamiltonien d'une particule dont le moment cinétique est proportionnel au dipôle magnétique, quand elle est plongée dans un champ magnétique uniforme (problème de la "précession de Larmor", dont la version quantique est d'une grande importance dans la compréhension de l'effet Zeeman). Les seules variables dynamiques importantes sont les composantes du moment cinétique. On peut écrire:

$$\frac{d\mathbf{L}}{dt} = \{H, \mathbf{L}\} = \{\mathbf{\Omega} \cdot \mathbf{L}, \mathbf{L}\} \tag{2.40}$$

En utilisant les relations de crochets de poisson entre les composantes de $\mathbf{L}$, on a, par exemple,

$$\frac{dL_x}{dt} = \Omega_y L_z - \Omega_z L_y \tag{2.41}$$

soit encore

$$\frac{d\mathbf{L}}{dt} = \mathbf{\Omega} \times \mathbf{L} \ . \tag{2.42}$$

On retrouve aisément que le moment cinétique précesse autour du vecteur fixe $\mathbf{\Omega}$ avec une pulsation Ω.

2.3 Action et hamiltonien

Nous allons revenir sur le principe variationnel. Nous allons écrire l'action en fonction du hamiltonien, et montrer qu'on peut retrouver les équations de Hamilton en écrivant la stationnarité de cette expression de l'action par rapport à certaines variations infinitésimales de la trajectoire. Ce paragraphe ne nous apprendra rien que nous ne sachions déjà, mais il est nécessaire à la cohérence de tout notre édifice.

Le lien entre fonction de Lagrange et fonction de Hamilton nous permet d'écrire l'action sous la forme:

$$S = \int_{t_1}^{t_2} \left(\sum_i p_i \dot{q}_i - H \right) \, dt \ , \tag{2.43}$$

avec les notations du premier chapitre.

Ecrivons maintenant, pour retrouver les équations de Hamilton, que l'action est stationnaire pour la trajectoire effectivement suivie. L'accroissement δS de l'action dans un accroissement infinitésimal de la trajectoire doit donc s'annuler. Nous considérerons donc deux trajectoires. La trajectoire de référence est la trajectoire effectivement suivie, définie par $q_i(t)$ et $p_i(t)$ (les variables naturelles dans une approche hamiltonienne sont les positions et les impulsions). L'autre trajectoire (infiniment proche) est définie à chaque instant par $q_i(t) + \delta q_i(t)$ et $p_i(t) + \delta p_i(t)$. Nous imposerons aux deux trajectoires de coïncider à l'instant initial et à l'instant final: $\delta q_i(1) = \delta q_i(2) = 0$. Il est naturel de considérer, dans une approche hamiltonienne, les p_i et les q_i comme des variables indépendantes. Nous n'imposerons donc aucune condition aux δp_i, ni aux extrémités de la trajectoire, ni à aucun instant. Il faut bien voir que nous pouvons ainsi considérer des trajectoires variées qui n'auraient pas de sens du point de vue de la simple cinématique: si les impulsions coïncident avec les quantités de mouvement ($m\mathbf{v}$), varier les vitesses indépendamment des positions implique que, sur la trajectoire variée, les vitesses puissent ne plus être égales aux dérivées des positions (avec une différence au premier ordre dans les petits accroissements). Ce n'est qu'au prix de cette liberté toute mathématique que nous pourrons retrouver l'équation de Hamilton qui contient en fait la définition des impulsions généralisées (de la vitesse dans le cas très simple que nous venons de mentionner).

Avec ces notations, l'accroissement de l'action s'écrit simplement:

$$\delta S = \int_{t_1}^{t_2} \left(\sum_i (\delta p_i \dot{q}_i + p_i \delta \dot{q}_i) - \delta H \right) \, dt \ . \tag{2.44}$$

L'accroissement de H entre la trajectoire de référence et la trajectoire variée s'écrit simplement:

$$\delta H = \sum_i \frac{\partial H}{\partial q_i}\delta q_i + \frac{\partial H}{\partial p_i}\delta p_i \tag{2.45}$$

On peut donc mettre δS sous la forme:

$$\delta S = \int_{t_1}^{t_2} \sum_i \delta p_i \left[\frac{dq_i}{dt} - \frac{\partial H}{\partial p_i}\right] dt + \int_{t_1}^{t_2} \sum_i \left[p_i \delta \dot{q}_i - \frac{\partial H}{\partial q_i}\delta q_i\right] dt \ . \tag{2.46}$$

Considérons plus particulièrement le premier terme de la deuxième intégrale. On peut l'intégrer par parties pour faire apparaître δq_i au lieu de $\delta \dot{q}_i$. Le terme tout intégré dans cette intégration par parties fait intervenir les accroissements δq_i aux instants t_1 et t_2. Il est donc identiquement nul. En regroupant alors les termes proportionnels aux accroissements des positions et des impulsions, on peut écrire:

$$\delta S = \int_{t_1}^{t_2} \sum_i \left[\frac{dq_i}{dt} - \frac{\partial H}{\partial p_i}\right] \delta p_i \ dt - \int_{t_1}^{t_2} \sum_i \left[\frac{dp_i}{dt} + \frac{\partial H}{\partial q_i}\right] \delta q_i \ dt \ . \tag{2.47}$$

δS ne s'annulera quels que soient les accroissements des positions et impulsions que si les termes entre crochets dans chaque intégrale s'annulent identiquement sur la trajectoire effectivement suivie. On montre bien ainsi que cette trajectoire obéit effectivement aux équations de Hamilton, que nous aurions parfaitement pu établir par ce raisonnement.

Il existe, dans le cadre de le formulation hamiltonienne, un autre principe variationnel, le principe de Maupertuis, qui permet de trouver la forme de la trajectoire, mais pas la loi horaire. Ce principe s'apparente de très près à celui de Fermat, qui permet de déterminer en optique la trajectoire des rayons lumineux. Nous n'aborderons pas ici le principe de Maupertuis, largement discuté dans les manuels[3].

2.4 Transformations canoniques

Nous abordons ici ce qui constitue sans doute l'intérêt essentiel de la formulation hamiltonienne. La complète symétrie entre les positions et impulsions va en effet nous permettre d'envisager des changements de variables mêlant ces deux types de quantités. Nous verrons qu'on peut ainsi rendre complètement triviale la dynamique d'un problème, en prenant par exemple comme nouvelles variables les conditions initiales. Bien sûr, la difficulté est de trouver le bon changement de variables. Nous n'aborderons pas, faute de place, les méthodes permettant de les déterminer. Nous profiterons également de ce paragraphe pour établir le théorème de Liouville, qui joue un rôle central en physique statistique et dans l'étude de la dynamique des systèmes complexes.

2.4.1 Principe

Revenons un instant au formalisme lagrangien. Les q_i et leurs vitesses $\dot{q}_i$ ne sont pas des variables indépendantes. Le seul type de changement de variable que l'on puisse envisager est donc de définir de nouvelles coordonnées généralisées Q_i, calculables à partir des q_i et du temps (la relation étant inversible). On peut alors écrire la fonction de Lagrange en fonction des Q_i, les équations de Lagrange correspondantes et résoudre ainsi le problème – en espérant qu'il soit plus simple en terme des nouvelles variables. Nous appellerons une telle transformation des coordonnées généralisées une "transformation ponctuelle". Nous allons maintenant voir qu'elle appartient à un cadre beaucoup plus général.

Dans le formalisme hamiltonien, les variables indépendantes sont les q_i et les p_i. Nous pouvons imaginer un changement de variables très général sous la forme $q_i \longrightarrow Q_i(q_i, p_i, t)$ et $p_i \longrightarrow P_i(q_i, p_i, t)$.

[3]On consultera en particulier le Goldstein.

Les définitions des nouvelles positions et des nouvelles impulsions font intervenir toutes les anciennes positions et impulsions. Ce changement de variables nous laisse donc toute liberté. Nous exigerons cependant qu'il soit inversible, pour que l'état du système soit déterminé de façon univoque par les nouvelles coordonnées, ce qui impose que les Q_i et les P_i soient indépendantes.

Pour que ce changement de variable soit utilisable dans un problème de mécanique, il nous faut cependant lui imposer une contrainte supplémentaire. Il faut en effet qu'il existe un nouveau hamiltonien $H'(Q_i, P_i, t)$ qui donne les équations de Hamilton pour les nouvelles variables. Nous dirons alors que la transformation est "canonique".

Supposons que H' existe bien. Le principe variationnel abordé au paragraphe précédent s'écrit:

$$\delta \int_{t_1}^{t_2} \left(\sum_i p_i \dot{q}_i - H \right) dt = 0 \tag{2.48}$$

en terme des anciennes variables et

$$\delta \int_{t_1}^{t_2} \left(\sum_i P_i \dot{Q}_i - H' \right) dt = 0 \tag{2.49}$$

en termes des nouvelles. Pour que ces deux principes variationnels donnent les mêmes équations du mouvement, il suffit que les deux quantités intégrées ne diffèrent que de la dérivée totale par rapport au temps d'une fonction F. En effet, cette différence ne contribue aux intégrales que par un terme de la forme $F(2) - F(1)$, qui ne varie pas quand on varie la trajectoire. On peut mettre cette condition sous la forme:

$$\sum_i p_i dq_i - H dt = \sum_i P_i dQ_i - H' dt + dF \ , \tag{2.50}$$

dF étant l'accroissement de F entre deux instants voisins. Il suffit donc, pour que la transformation soit canonique, qu'il existe une fonction F telle que:

$$dF = \sum_i p_i dq_i - \sum_i P_i dQ_i + (H' - H) dt \tag{2.51}$$

ou encore:

$$\frac{\partial F}{\partial q_i} = p_i, \qquad -\frac{\partial F}{\partial Q_i} = P_i, \qquad H' = H + \frac{\partial F}{\partial t} \tag{2.52}$$

Si effectivement la donnée d'une fonction $F(q_i, Q_i, t)$ permet, en écrivant les relations aux dérivées partielles précédentes, de déterminer complètement la transformation, celle-ci sera une transformation canonique et le nouveau hamiltonien sera connu.

Pour montrer que la donnée de F détermine complètement la transformation, considérons les n équations $\partial F(q_i, Q_i, t)/\partial q_i = p_i$. On peut, au moins formellement et sauf cas pathologique, les résoudre en termes des n "inconnues" Q_i et trouver ainsi les n fonctions $Q_i(q_i, p_i, t)$. On peut alors calculer facilement, en fonction des q_i, Q_i et de t les n dérivées partielles $-\partial F/\partial Q_i$ qui donnent les nouvelles impulsions $P_i(q_i, Q_i, t)$. En reportant les expressions des Q_i, on trouve alors les $P_i(q_i, p_i, t)$, ce qui achève de déterminer complètement la transformation. Le nouveau hamiltonien H' peut alors être calculé et on peut écrire les équations de Hamilton en termes des nouvelles variables.

En résumé, la donnée d'une fonction $F(q_i, Q_i, t)$ détermine en général de façon univoque une transformation généralisée et assure que cette transformation soit canonique. Enfin, la donnée de F permet d'exprimer le nouveau hamiltonien. Pour toutes ces raisons F s'appelle la fonction génératrice de la transformation[4].

[4]Nous avons montré qu'à toute fonction génératrice correspond une transformation (sauf cas pathologique, certaines fonctions ne définissant pas une transformation univoque – en particulier les constantes). Nous n'avons pas établi la réciproque, que nous admettrons.

Nous avons choisi ici, de façon naturelle, de définir F par les variables q_i et Q_i. On peut imaginer de définir une transformation canonique par un autre couple de variables. F étant donnée, considérons en effet la fonction Φ définie par:

$$\Phi = F + \sum_i P_i Q_i \ , \tag{2.53}$$

et résultant donc d'une transformation de Legendre sur la fonction F. ϕ est donc a priori une fonction naturelle des q_i, P_i. On peut s'en convaincre en écrivant sa différentielle:

$$d\phi = \sum_i p_i dq_i + \sum_i Q_i dP_i + (H' - H)dt \ . \tag{2.54}$$

Les dérivées partielles de ϕ sont donc:

$$\frac{\partial \phi}{\partial q_i} = p_i, \qquad \frac{\partial \phi}{\partial P_i} = Q_i, \qquad H' = H + \frac{\partial \phi}{\partial t} \ . \tag{2.55}$$

Comme dans le cas précédent, on peut montrer aisément que la donnée de ϕ et les relations aux dérivées partielles ci–dessus déterminent complètement la transformation canonique. Pour cela, on résout les n équations $\partial\phi/\partial q_i = p_i$ en termes des $P_i(q_i, p_i, t)$. En reportant ces expressions dans les n autres relations, on achève de déterminer la transformation en obtenant les Q_i.

Nous laissons au lecteur le soin de montrer qu'il existe encore deux expressions possibles pour une transformation canonique. L'une fait intervenir $\Psi(p_i, Q_i, t) = F - \sum_i q_i p_i$ et les relations:

$$\frac{\partial \Psi}{\partial p_i} = -q_i, \qquad \frac{\partial \Psi}{\partial Q_i} = -P_i \qquad H' = H + \frac{\partial \Psi}{\partial t} \ . \tag{2.56}$$

L'autre utilise la fonction $\Xi(p_i, P_i, t) = \phi - \sum_i Q_i P_i$ et les relations:

$$\frac{\partial \Xi}{\partial p_i} = -q_i, \qquad \frac{\partial \Xi}{\partial P_i} = Q_i \qquad H' = H + \frac{\partial \Xi}{\partial t} \ . \tag{2.57}$$

La donnée au choix de l'une de ces quatre fonctions détermine donc complètement une transformation canonique. Si la fonction génératrice ne fait pas intervenir explicitement le temps, les fonctions de Hamilton coïncident dans les anciennes et les nouvelles représentations (il existe en fait un lien très profond entre les transformations canoniques et les changements de représentation ou les transformations unitaires en mécanique quantique). Le choix immense des fonctions génératrices possibles donne une idée de la puissance de la méthode.

2.4.2 Transformations canoniques et crochets de Poisson

En fait la propriété essentielle des transformations canoniques est qu'elles conservent les crochets de Poisson:

$$\{f, g\}_{p,q} = \{f, g\}_{P,Q} \ , \tag{2.58}$$

où

$$\{f, g\}_{p,q} = \sum_i \frac{\partial f(p_i, q_i, t)}{\partial p_i} \frac{\partial g}{\partial q_i} - \frac{\partial f}{\partial q_i} \frac{\partial g}{\partial p_i} \ , \tag{2.59}$$

et

$$\{f, g\}_{P,Q} = \sum_i \frac{\partial f(P_i, Q_i, t)}{\partial P_i} \frac{\partial g}{\partial Q_i} - \frac{\partial f}{\partial Q_i} \frac{\partial g}{\partial P_i} \ . \tag{2.60}$$

On peut bien sûr démontrer cette relation en utilisant la fonction génératrice de la transformation. Il existe cependant une démonstration beaucoup plus intuitive. Limitons nous pour cela au cas où f et g ne dépendent pas explicitement du temps et au cas où la fonction génératrice de la transformation ne dépend pas non plus du temps. Il doit exister un problème de mécanique décrit par les q_i, p_i dont la

fonction $g(p_i, q_i)$ serait le hamiltonien. Le crochet de Poisson $\{f, g\}_{p,q}$ apparaît alors comme la dérivée temporelle de la fonction f le long d'un trajectoire solution de ce problème fictif. La fonction $g(P_i, Q_i)$ est le hamiltonien H' du même problème dans les coordonnées transformées. Le crochet de Poisson $\{f, g\}_{P,Q}$ est la dérivée temporelle de f dans cette représentation. Mais df/dt doit être indépendante de la représentation, ce qui implique l'égalité des crochets de Poisson. Cette relation, établie ici dans un cas un peu particulier, est en fait générale. Elle permet, par exemple, de calculer facilement les crochets de poisson des nouvelles variables dans la représentation des anciennes, ce que nous laisserons au lecteur à titre d'exercice.

2.4.3 Exemples de transformations canoniques

Nous donnons ici, à titre d'illustration, quelques fonctions génératrices définissant des transformations canoniques particulièrement simples. Nous montrerons aussi comment les transformations canoniques permettent de rendre triviale la solution d'un problème de mécanique.

- $\Phi(q_i, P_i, t) = \sum_i q_i P_i$. Le temps n'intervenant pas dans la fonction génératrice, nous avons bien sûr $H' = H$. En appliquant les relations aux dérivées partielles, nous trouvons sans difficultés $p_i = P_i$ et $Q_i = q_i$. Cette fonction génératrice définit donc la transformation unité, ce qui ne présente guère d'interêt.

- $\Phi(q_i, P_i, t) = \sum_i \phi_i(q_j, t) P_i$. Là encore, on trouve sans difficultés $Q_i = \phi_i(q_j, t)$. Cette fonction génère donc l'ensemble des transformations ponctuelles, éventuellement dépendantes du temps, qui définissent les nouvelles cordonnées en fonction seulement des anciennes. L'avantage de l'approche en termes de transformations canoniques est que l'application des autres relations aux dérivées partielles nous donne $p_i = \sum_k (\partial\phi_k/\partial q_i) P_i$, système qui permet de déterminer les nouvelles impulsions en fonction des anciennes et des coordonnées. Notons que, si les ϕ_i dépendent du temps, les deux hamiltoniens H et H' peuvent différer.

- $F = \sum_i (q_i Q_i)$. On trouve alors immédiatement $p_i = Q_i$ et $P_i = -q_i$. A un signe près, cette fonction réalise l'échange des coordonnées et des impulsions, illustrant le rôle très symétrique que jouent ces notions dans l'approche hamiltonienne.

Nous allons montrer maintenant comment une transformation canonique bien choisie permet de rendre complètement triviale la dynamique d'un système. L'idée est de rendre le nouveau hamiltonien cyclique dans les nouvelles coordonnées. Les nouvelles impulsions sont alors constantes et la dynamique des nouvelles coordonnées se résume à une évolution linéaire dans le temps. La difficulté dans ce genre d'approche, à la base de nombreuses méthodes de résolution de problèmes de mécanique, est bien sûr d'exhiber la transformation canonique convenable, ce qui n'est pas toujours possible. Nous considérerons dans ce paragraphe le problème trivial d'un oscillateur harmonique à une dimension.

Le hamiltonien H s'exprime simplement en fonction de la coordonnée q et de l'impulsion conjuguée p par:

$$H = \frac{p^2}{2m} + \frac{m\omega^2}{2} q^2 \ , \tag{2.61}$$

somme des énergies cinétiques et potentielles. On considère alors la transformation canonique générée par la fonction

$$F(q, Q, t) = \frac{1}{2} m\omega q^2 \cot Q \tag{2.62}$$

(le choix d'une telle transformation n'est guère possible si on ne connaît déjà la solution du problème). Cette fonction ne dépendant pas explicitement du temps, le nouveau et l'ancien hamiltonien coïncident. A partir de cette fonction génératrice, on trouve

$$p = \frac{\partial F}{\partial q} = m\omega q \cot Q \tag{2.63}$$

et

$$P = -\frac{\partial F}{\partial Q} = \frac{m\omega q^2}{2\sin^2 Q} \tag{2.64}$$

On peut extraire de ces deux relations:

$$q = \sqrt{\frac{2P}{m\omega}}\sin Q \ , \qquad p = \sqrt{2m\omega P}\cos Q \ . \tag{2.65}$$

Le nouveau hamiltonien peut alors être exprimé, comme il se doit, en termes des nouvelles variables:

$$H' = \frac{p^2}{2m} + \frac{m\omega^2}{2}q^2 = \omega P \ . \tag{2.66}$$

Comme nous l'espérions, ce nouveau hamiltonien est cyclique dans la nouvelle coordonnée Q. L'impulsion conjuguée, P, est donc une constante. Comme l'énergie mécanique totale est conservée dans ce problème et coïncide avec la fonction de Hamilton, cette constante vaut simplement:

$$P = \frac{\mathcal{E}}{\omega} \ . \tag{2.67}$$

Elle est donc homogène à une action (produit d'une énergie par un temps). P s'appelle donc variable d'action. On montre que, dans tout problème unidimensionnel, on peut trouver une variable d'action conservée dans l'évolution. L'équation de Hamilton pour Q, $\dot{Q} = \partial H'/\partial P = \omega$, donne simplement:

$$Q = \omega t + \phi \ . \tag{2.68}$$

Q, évoluant linéairement avec le temps, se nomme variable d'angle. Là encore, dans tout problème unidimensionnel, il existe une variable d'angle conjuguée de la variable d'action. La solution explicite du problème est donc donnée en fonction de deux constantes arbitraires, comme il se doit, l'énergie mécanique totale $\mathcal{E}$ et la phase ϕ, valeur initiale de la nouvelle coordonnée. En utilisant alors la transformation, on peut exprimer la solution en termes des variables initiales, et on trouve:

$$q = \sqrt{\frac{2\mathcal{E}}{m\omega^2}}\sin(\omega t + \phi) \ , \tag{2.69}$$

ce qui n'est pas vraiment un résultat inattendu!

L'apparente simplicité de cette approche, dans ce cas trivial, ne doit pas faire oublier que la grande difficulté est d'exhiber la fonction génératrice adaptée. Nous ne pourrons aborder ici les méthodes variées de résolution fondées sur les transformations canoniques. Le lecteur pourra en trouver une description détaillée dans le Goldstein.

2.4.4 Transformations canoniques et espace des phases

L'état mécanique du système est complètement décrit par la donnée des n q_i et des n p_i. Autrement dit, le système est décrit comme un point dans un espace à $2n$ dimensions que l'on appelle *espace des phases*. Cet espace joue un très grand rôle en physique statistique. L'entropie, par exemple, peut être définie comme le logarithme du nombre de configurations accessibles au système. Compter ce nombre de configurations, c'est compter la surface de l'espace des phases correspondant à un petit intervalle d'énergie. L'espace des phases joue également un rôle très important dans l'étude de la dynamique complexe des système (chaos classique, par exemple). Un point d'équilibre stable correspond à un point dans l'espace des phases, un mouvement périodique à une trajectoire fermée simple, un mouvement chaotique à une trajectoire complexe parcourant rapidement tout le domaine accessible.

Les transformations canoniques transforment un espace des phases en un autre. La géométrie de cette transformation n'est pas complètement arbitraire, en raison des contraintes imposées aux

transformations canoniques. La propriété essentielle est qu'une transformation canonique conserve le volume dans l'espace des phases. Si on considère un domaine V de l'espace des phases des "anciennes coordonnées", il lui correspond un domaine V' dans le nouvel espace. Pour des raisons évidentes de continuité et dérivabilité des transformations canoniques, V' est un fermé connexe si V l'est. On peut calculer le volume $\mathcal{V}$ du domaine V comme:

$$\mathcal{V} = \int_V dq_1 \cdots dq_n dp_1 \cdots dp_n \ , \tag{2.70}$$

et calculer de même le volume $\mathcal{V}'$ du domaine V'.

Pour une transformation canonique, $\mathcal{V} = \mathcal{V}'$.

Pour les mathématiciens cette propriété découle de façon immédiate de la structure symplectique de la transformation canonique. Nous allons établir cette propriété de façon moins directe, mais peut être plus accessible. Rappelons d'abord que:

$$\int_{V'} dQ_1 \cdots dQ_n dP_1 \cdots dP_n = \int_V |J| \ dq_1 \cdots dq_n dp_1 \cdots dp_n \ , \tag{2.71}$$

où J est le Jacobien du changement de variable, déterminant formé avec toutes les dérivées partielles des nouvelles variables par rapport aux anciennes:

$$J = \begin{vmatrix} \dfrac{\partial Q_1}{\partial q_1} & \cdots & \dfrac{\partial P_n}{\partial q_1} \\ \vdots & \ddots & \vdots \\ \dfrac{\partial Q_1}{\partial p_n} & \cdots & \dfrac{\partial P_n}{\partial p_n} \end{vmatrix} \ , \tag{2.72}$$

que nous noterons également

$$J = \frac{\partial(Q_1, \ldots, P_n)}{\partial(q_1, \ldots, p_n)} \ . \tag{2.73}$$

Ces notations deviennent tout à fait triviales en dimension 1 et coïncident alors avec les changements de variables standard dans les intégrales. Si nous prouvons que le Jacobien de toute transformation canonique est 1, nous aurons établi la proposition cherchée.

Pour cela, nous aurons besoin de deux propriétés des Jacobiens. D'abord:

$$\frac{\partial(Q_1, \ldots, P_n)}{\partial(q_1, \ldots, p_n)} = \frac{\dfrac{\partial(Q_1, \ldots, Q_n, P_1, \ldots, P_n)}{\partial(q_1, \ldots q_n, P_1, \ldots, P_n)}}{\dfrac{\partial(q_1, \ldots, q_n, p_1, \ldots, p_n)}{\partial(q_1, \ldots q_n, P_1, \ldots, P_n)}} \ . \tag{2.74}$$

En fait, comme les dérivées partielles ordinaires, les produits et rapports de Jacobiens peuvent se simplifier comme des fractions[5]. De plus:

$$\frac{\partial(Q_1, \ldots, Q_n, P_1, \ldots, P_n)}{\partial(q_1, \ldots, q_n, P_1, \ldots, P_n)} = \frac{\partial(Q_1, \ldots, Q_n)}{\partial(q_1, \ldots, q_n)} \ . \tag{2.75}$$

On peut donc retirer d'un Jacobien les variables qui apparaissent au "numérateur" et au "dénominateur". Le changement de variables considéré laisse en effet invariantes ces quantités.

En utilisant successivement ces deux propriétés, on met le Jacobien de la transformation canonique sous la forme:

$$J = \frac{\dfrac{\partial(Q_1, \ldots, Q_n)}{\partial(q_1, \ldots, q_n)}}{\dfrac{\partial(p_1, \ldots, p_n)}{\partial(P_1, \ldots, P_n)}} \ . \tag{2.76}$$

[5]Nous supposons bien sûr que tous les Jacobiens écrits dans ces équations ont un sens, et en particulier que toutes les transformations "tronquées" sont inversibles, ce qui n'est pas vrai en toute généralité.

La transformation étant canonique, elle est engendrée par une fonction $\Phi(q_i, P_i)$ telle que $\partial\Phi/\partial q_i = p_i$ et $\partial\Phi/\partial P_i = Q_i$. On a donc immédiatement:

$$\frac{\partial Q_i}{\partial q_j} = \frac{\partial^2 \Phi}{\partial q_j \partial P_i} \tag{2.77}$$

et

$$\frac{\partial p_i}{\partial P_j} = \frac{\partial^2 \phi}{\partial P_j \partial q_i} \ . \tag{2.78}$$

Les dérivées secondes croisées étant égales, toutes les dérivées partielles apparaissant dans le développement du déterminant au numérateur de (2.76) sont égales, terme à terme, à celles apparaissant dans le développement du dénominateur. Cela établit que le Jacobien d'une transformation canonique est de module égal à un et l'invariance du volume dans l'espace des phases.

2.4.5 Transformation générée par l'action et théorème de Liouville

Nous considérerons dans ce paragraphe une transformation canonique très particulière qui fait se correspondre deux états du système à deux instants différents. Considérons en effet un système ayant un lagrangien indépendant du temps (l'énergie totale est donc conservée) dont la dynamique est décrite par les $p_i(t)$ et les $q_i(t)$ et considérons le changement de variables défini par:

$$p_i(t) \quad \longrightarrow \quad P_i(t) = p_i(t+T) \tag{2.79}$$

$$q_i(t) \quad \longrightarrow \quad Q_i(t) = q_i(t+T) \ , \tag{2.80}$$

où T est une durée fixe. Cette correspondance entre états du système à des instants différents est évidemment une transformation canonique, puisque les nouvelles variables obéissent aux mêmes équations du mouvement que les anciennes. Nous allons voir que la fonction génératrice de cette transformation n'est autre que l'action. Considérons en effet la trajectoire du système entre les instants t et $t+T$ et une trajectoire infiniment voisine obtenue en modifiant les coordonnées aux points de départ et d'arrivée de quantités infinitésimales. Avec les résultats du chapitre précédent, nous pouvons écrire la variation de l'action entre ces deux trajectoires comme:

$$dS = \sum_i p_i(t+T) dq_i(t+T) - \sum_i p_i(t) dq_i(t) \tag{2.81}$$

ou encore

$$dS = \sum_i P_i(t) dQ_i(t) - \sum_i p_i(t) dq_i(t) \ . \tag{2.82}$$

Cette expression de la différentielle de l'action, que nous pouvons considérer comme une fonction des coordonnées des points de départ et d'arrivée: $S(q_i, Q_i)$, prouve que:

$$\frac{\partial S}{\partial Q_i} = P_i \qquad \frac{\partial S}{\partial q_i} = -p_i \ . \tag{2.83}$$

S est donc bien fonction génératrice de la transformation des q_i en Q_i. Cette transformation joue un rôle central dans la méthode de Hamilton–Jacobi, essentielle pour la résolution de problèmes complexes, et conduisant à la notion importante de séparabilité des variables. Nous ne disposons pas d'un espace suffisant pour traiter convenablement cette méthode. Nous donnerons donc une seule application de la transformation engendrée par l'action.

Considérons un domaine V de l'espace des phases du système. On peut considérer l'ensemble des trajectoire originaires d'un point situé à l'instant t à l'intérieur de ce domaine. Par continuité, ces trajectoires correspondent à l'instant $t+T$ à des points situés dans un nouveau domaine V' de l'espace des phases. Comme la transformation faisant se correspondre les instants t et $t+T$ est canonique,

l'étendue du domaine V' est égale à celle du domaine V. Cette propriété constitue le théorème de Liouville:

Le volume du domaine occupé dans l'espace des phases par un ensemble de trajectoires se conserve au cours du temps.

Ce théorème joue évidemment un rôle important en mécanique statistique. Il prédit, par exemple, la conservation de l'entropie dans une évolution hamiltonienne. Notons que ce théorème ne tient pas en présence de dissipation. N'importe quelle condition initiale conduit en effet à un état de repos où les coordonnées n'évoluent plus. Soulignons aussi le lien entre ce théorème et le théorème de conservation de l'étendue en optique. L'étendue joue le rôle du volume dans un espace des phases {position des rayons/angle }.

Ce paragraphe clôt notre exposé de mécanique analytique. Nous n'avons pas, de loin, donné un exposé exhaustif de ce sujet. Les grands domaines que nous n'aborderons pas sont les méthodes de résolution, telles que la méthode de Hamilton Jacobi. Nous ne dirons rien, non plus, des méthodes de perturbations classiques, si utiles en astronomie. Nous évoquerons brièvement dans la prochaine partie les extensions du formalisme lagrangien à des coordonnées continues (en un mot à des champs), mais sans épuiser non plus ce très vaste sujet. Finalement, nous ne saurions trop recommander au lecteur de se rapporter aux manuels de mécanique quantique pour explorer les liens très profonds entre mécanique quantique et mécanique analytique. Si une présentation de la mécanique quantique à partir de la dynamique classique et de sa quantification canonique n'est pas à recommander pour une première approche de la mécanique quantique, elle est extrêmement enrichissante à un niveau plus avancé.

Appendice 1

Modèle de Bohr

Nous traiterons dans cet appendice du premier modèle réaliste de structure atomique, celui de Bohr en 1913. Il nous sera en effet nécessaire à plusieurs endroits du cours pour des discussions qualitatives du rayonnement atomique. Comme ce modèle fait explicitement référence aux concepts d'action et a donné lieu, au cours de son évolution, à des développements élégants de mécanique classique, il a tout naturellement sa place dans ce cours de mécanique analytique. Nous commencerons par un bref rappel de la situation historique au moment de la formulation du modèle de Bohr. Nous l'exposerons ensuite dans un deuxième paragraphe. Dans un dernier paragraphe, nous préciserons ses limitations et les tentatives, menées entre autres par Bohr et Sommerfeld, pour raffiner le modèle et l'appliquer à d'autres atomes que l'hydrogène. Enfin, nous rappellerons brièvement, pour mémoire, les résultats quantiques rigoureux.

1.1 Un peu d'histoire

Les difficultés de la mécanique classique et les débuts de la mécanique quantique, au tournant du siècle, sont dues à essentiellement deux problèmes: le rayonnement du corps noir et la structure des spectres d'émission ou d'absorption des vapeurs.

Quand on traite en thermodynamique classique le rayonnement d'un corps complètement absorbant en équilibre thermique, on trouve la célèbre loi de Rayleigh Jeans. La densité de puissance spectrale du rayonnement (quantité d'énergie par unité de volume et de fréquence) est proportionnelle au carré de la fréquence. La quantité totale d'énergie électromagnétique contenue dans un corps en équilibre devrait donc être gravement infinie. Ce n'est bien sûr pas le cas et les données expérimentales, relativement précises à la fin du siècle dernier, donnaient un spectre décroissant rapidement à haute fréquence. Pour expliquer ce spectre, Planck introduisit, en 1900, une hypothèse de quantification. Les échanges d'énergie entre matière et rayonnement ne peuvent se faire, à une fréquence donnée, que par multiples entiers d'une quantité fondamentale, proportionnelle à la fréquence, selon la fameuse relation $E = h\nu$. La constante de Planck, h, une fois ajustée aux données expérimentales, l'accord entre les spectres calculés et les spectres expérimentaux se révélait excellent. En fait, Planck considérait cette hypothèse comme heuristique et doutait de sa signification physique. Ce n'est qu'avec Einstein, 5 ans plus tard, que l'idée de quantification de l'énergie électromagnétique fit une avancée notable avec l'introduction de quanta lumineux. Ceux-ci, qu'on devait appeler plus tard photons, sont nécessaires pour analyser, au-delà des valeurs moyennes, les fluctuations du rayonnement. On peut alors interpréter convenablement les propriétés de l'effet photoélectrique, ce qui valut son Nobel à Einstein (la relativité générale paraissait peut-être trop audacieuse pour être couronnée).

L'histoire de la spectroscopie est pour sa part très riche. Dès 1802, Wollaston, avec un spectrographe à prisme, observait des bandes sombres bien résolues dans le spectre solaire, bandes qui ne

sont pas présentes dans le spectre des corps chauffés[1]. C'est sans doute là la première observation d'un spectre d'absorption atomique. Wollaston décrit en particulier une bande très intense dans le jaune qui devait être la raie principale du sodium. En 1817, Fraunhofer, encore très jeune, raffine ces mesures et observe les raies "de Balmer" du spectre de l'hydrogène (l'appellation Balmer est beaucoup plus récente comme on le verra). Il invente peu après le réseau de diffraction et l'histoire de la spectroscopie de précision commence. Il mesure avec grand soin (et avec une précision de l'ordre de 10^{-4}) la position des raies de Balmer. La nature de ces raies reste toutefois très controversée. En parallèle avec ces études du spectre solaire, l'étude du spectre d'émission de décharges dans les gaz se poursuit. Masson, en 1851, franchit un pas important en montrant que certaines raies peuvent être attribuées sans ambiguïté à la présence d'hydrogène. C'est la première identification fiable d'un élément avec un spectre de raies. Un nouveau pas important est franchi avec Ångström, qui remarque et explique la coïncidence des raies d'émission de l'hydrogène avec les raies d'absorption dans le spectre solaire. Il conclut que tout élément peut aussi bien absorber ou émettre de la lumière sur une de ses fréquences propres. Des atlas détaillés du spectre solaire sont publiés ensuite par Ångström, Rowland et Huggins (c'est sur les travaux de ce dernier que Balmer s'est, semble-t-il, appuyé).

On a cherché dans le même temps, en manipulant les données spectrales, à dégager des lois auxquelles pourraient obéir les fréquences des raies spectrales, en particulier pour l'hydrogène dont le spectre est simple. On pensa ainsi, pendant un temps, que les différentes fréquences émises par l'hydrogène pourraient être des harmoniques d'une fréquence fondamentale (travaux de Stoney en 1871). Les coïncidences numériques supportant cette approche s'évanouirent rapidement avec les progrès de l'instrumentation et en particulier ceux des spectrographes à réseau.

Un progrès important fut accompli par Balmer en 1885. Déjà âgé, simple instituteur, il manipule les données sur le spectre de l'hydrogène. Il observe que les longueurs d'onde des raies visibles du spectre de l'hydrogène sont proportionnelles à des fractions rationnelles simples faisant intervenir les carrés des nombres entiers, sous la forme $m^2/(m^2 - n^2)$, avec $n = 2$. Ce travail remarquable fut complété par celui de Rydberg, en 1889, qui remarque qu'il vaut mieux considérer les nombres d'ondes (inverses de longueurs d'onde). Les nombres d'ondes des raies de nombreux éléments s'obtiennent en effet simplement comme différences de termes en R/n^2 où n est entier et où R est la maintenant célèbre constante de Rydberg. Le principe de combinaison de Ritz, formulé en 1908, généralise ce travail. La découverte de nouvelles séries de raies de l'hydrogène s'accordant avec les formules de Balmer ($n = 3$ par Paschen en 1908, $n = 1$ (ultraviolet lointain) par Lyman en 1916, $n = 4$ par Brackett en 1922, $n = 5$ par Pfund en 1924, $n = 6$ par Humphrey en 1953) apporta au cours du temps des confirmations remarquables de la formule de Balmer. Bien sûr, la découverte par Michelson et Morley (encore eux) de sous–structures dans les raies de l'hydrogène (nous dirions aujourd'hui de la structure fine) complique un peu le tableau, mais le succès des formules de Balmer ou de Rydberg demeure.

S'il existe une relation aussi simple entre les fréquences, on doit chercher un modèle physique qui les prédise correctement. Le premier modèle "réaliste" de la structure de l'atome d'hydrogène est dû à Thomson, découvreur de l'électron en 1897 au Cavendish Laboratory, fondé par Maxwell environ 20 ans avant (en fait la découverte de l'électron pourrait aussi bien être attribuée à Wiechert, qui travaillait indépendamment). Comme on savait que la matière contient des électrons, Thomson imagine un atome constitué d'une gelée, uniformément chargée positivement, de forme sphérique, dans laquelle se déplacent des électrons indépendants. Le champ électrique produit par la gelée étant proportionnel à la distance, les électrons sont élastiquement liés et effectuent un mouvement harmonique de fréquence donnée. Nous utiliserons assez largement ce modèle très simple dit du "plum–pudding" dans le chapitre sur le rayonnement des sources atomiques. Nous verrons qu'il prédit correctement de nombreux ordres de grandeur.

Ce modèle, en dépit de certains succès, dut être abandonné après les expériences de déviation de

[1]Cette brève histoire de la spectroscopie s'inspire d'un article de G.W. Series, dans *The Hydrogen Atom*, Bassani et al. éditeurs, Springer, 1989.

particules α dans des feuilles d'or. En interprétant ces résultats, Rutherford, lui aussi directeur du Cavendish, fut conduit en 1910 à admettre la présence dans la matière de charges positives extrêmement localisées. Il fallait donc renoncer au modèle du plum–pudding et venir à un modèle planétaire de la structure atomique, avec des électrons orbitant sous l'influence de la force de Coulomb autour d'un noyau pratiquement ponctuel. Bien évidemment, un tel système rayonnerait ou absorberait à la fréquence de révolution de l'électron.

Ce modèle présente des difficultés sérieuses. La première est que rien a priori ne fixe les paramètres de l'orbite et donc la fréquence d'émission. On pourrait donc s'attendre à voir les atomes rayonner ou absorber des longueurs d'onde arbitraires. De façon plus grave, ce modèle est manifestement instable. En rayonnant, l'électron en mouvement perd de l'énergie et le rayon de son orbite diminue (nous ferons le calcul explicitement dans la quatrième partie). La fréquence du mouvement augmentant, la perte d'énergie et la chute vers le noyau deviennent de plus en plus rapide. En quelques dizaines de picosecondes, tous les électrons de l'univers auraient dû tomber sur leur noyau en émettant un bref flash de radiation ultraviolette. Cette catastrophe ultraviolette ne s'étant pas produite encore, il faut y voir un grave défaut du modèle.

1.2 Modèle de Bohr

L'attitude de Bohr est tout à la fois pragmatique et extrêmement audacieuse. Puisque rien ne peut expliquer la stabilité et le caractère discret de la structure atomique, c'est qu'il faut introduire dans le modèle une condition supplémentaire de "quantification". Dans cette démarche, Bohr était sans doute guidé par les résultats de Planck. Puisque la constante de Planck décrit la quantification pour le rayonnement électromagnétique, il est assez naturel de tenter de l'utiliser pour la structure atomique. Cette constante ayant la dimension d'une action, il est naturel aussi de quantifier l'action de l'électron sur sa trajectoire[2].

Nous poserons donc que l'action, calculée sur une orbite, est un multiple entier (évidemment non nul) de la constante de Planck:

$$S = nh \ . \tag{1.1}$$

Nous allons utiliser cette condition pour déterminer l'énergie de l'orbite, c'est à dire le terme spectral de Ritz qui lui est associé. Les fréquences des différentes transitions s'obtiendront comme des différences de ces termes spectraux.

Le gradient de l'action par rapport à l'extrémité de la trajectoire étant l'impulsion, cette quantité s'écrit évidemment:

$$S = \int \mathbf{p} \cdot d\mathbf{r} \ , \tag{1.2}$$

$\mathbf{r}$ étant la position de l'électron sur son orbite elliptique. L'action dépend donc de l'énergie de l'orbite, fixée par le demi grand axe de l'ellipse, mais aussi de l'excentricité de celle-ci. Pour fixer les paramètres de l'orbite, il faut imposer une condition supplémentaire. Bohr considère donc seulement des orbites circulaires. Une telle limitation, très arbitraire, n'est justifiée que par son succès. Sur une telle orbite les modules de la vitesse $\mathbf{v}$ et de $\mathbf{r}$ sont constants et on a évidemment:

$$S = 2\pi m r v = 2\pi L = nh \tag{1.3}$$

où m est la masse de l'électron et L la norme du moment angulaire. Notons que nous devrions utiliser ici en toute rigueur la masse réduite de l'électron qui tient compte de l'effet d'entraînement du noyau. Cet effet et même sa variation d'un isotope de l'hydrogène à l'autre sont parfaitement mesurables. Pour simplifier, nous considérerons dans la suite la masse du proton comme infinie. Les constantes que nous définirons sont alors exprimées en fonction de la vraie masse de l'électron et il faudrait appliquer des facteurs correctifs aux formules pour tenir compte de l'effet d'entraînement.

[2]Nous ne reproduisons pas ici les arguments originels de Bohr, un peu moins accessibles.

La condition de quantification de Bohr s'écrit donc aussi:

$$L = n\hbar \ , \tag{1.4}$$

où $\hbar = h/2\pi$. C'est sous cette forme qu'elle est le plus souvent écrite.

En 1923, De Broglie introduit une onde associée à toute particule quantique dont la longueur d'onde est donnée par la célèbre relation:

$$\lambda = h/p \ . \tag{1.5}$$

On constatera sans peine que la relation de quantification de Bohr est équivalente à postuler que le périmètre de l'orbite est égal à un nombre entier de longueurs d'onde de de Broglie. Cette condition "d'onde stationnaire" est très suggestive mais ne pourra être employée de façon quantitative avant le développement de l'équation de Schrödinger en 1926.

En notant e le module de la charge de l'électron et en écrivant l'équilibre de l'électron sur sa trajectoire circulaire, on montre immédiatement que

$$v = \sqrt{\frac{e^2}{4\pi\epsilon_0 rm}} \ . \tag{1.6}$$

En reportant cette expression dans la condition de quantification, on trouve le rayon de l'orbite:

$$r = a_0 n^2 \ , \tag{1.7}$$

où a_0, rayon de Bohr, est défini par:

$$a_0 = \frac{4\pi\epsilon_0\hbar^2}{me^2} \ . \tag{1.8}$$

Numériquement, le rayon de Bohr, qui est le rayon de l'état fondamental de l'hydrogène, vaut 0.053 nm.

Il est physiquement intéressant de comparer le rayon de Bohr à une longueur caractéristique formée avec les paramètres de l'électron et la constante de Planck. Il s'agit de la longueur d'onde de Compton de l'électron:

$$\lambda_c = \frac{h}{mc} \tag{1.9}$$

(on se référera au chapitre sur la relativité restreinte pour une description détaillée de l'effet Compton, collision d'un photon énergétique et d'un électron). On peut écrire:

$$a_0 = \frac{\lambda_c}{2\pi\alpha} \ , \tag{1.10}$$

où:

$$\alpha = \frac{e^2}{4\pi\epsilon_0\hbar c} \tag{1.11}$$

(c est la vitesse de la lumière dans le vide). Cette constante, sans dimensions, numériquement égale à 1/137, joue un rôle essentiel dans le modèle de Bohr et au delà dans toute l'électrodynamique quantique. Si, pour des raisons purement historiques, elle est appelée "constante de structure fine", elle mesure en fait la "force" de l'interaction électromagnétique. C'est en effet la seule constante sans dimension formée avec les paramètres de l'électromagnétisme (charge de l'électron et vitesse de la lumière) et la constante de Planck. Pratiquement tous les résultats de l'électrodynamique quantique peuvent se mettre sous la forme d'une fonction simple de cette constante ou d'un développement en ses puissances. Le modèle de Bohr ne fera pas exception à la règle.

Il est maintenant trivial de calculer l'énergie de l'électron sur son orbite et donc le terme spectral de Ritz. On a

$$E = -\frac{R}{n^2} \tag{1.12}$$

(cette énergie, correspondant à un état lié, est évidemment négative). On retrouve bien les termes spectraux en $1/n^2$ de la formule de Balmer. La constante R, qui n'est autre que la constante de Rydberg, peut s'écrire:

$$R = \frac{e^2}{8\pi\epsilon_0 a_0} \ . \tag{1.13}$$

On vérifiera sans peine qu'elle s'écrit aussi, en termes de la constante de structure fine et de l'énergie relativiste de masse de l'électron, mc^2, comme:

$$R = mc^2 \frac{\alpha^2}{2} \ . \tag{1.14}$$

On vérifiera sans peine que la valeur numérique est de 13.6 eV.

Nous avons ici calculé les fréquences de transition comme si la masse du proton était infinie. Pour une masse finie, on trouve encore bien sûr une loi en $1/n^2$, avec une constante de Rydberg légèrement modifiée $R_M = R/(1+m/M)$ où M est la masse de noyau (1836 m pour l'hydrogène). Les fréquences des raies émises s'interprètent alors simplement. L'atome peut effectuer une transition entre deux niveaux quantiques en émettant ou en absorbant un photon ayant une énergie égale à la différence des énergies du niveau initial et du niveau final. La fréquence de la transition entre les niveaux n et m est donc de la forme $(R/h)(1/n^2 - 1/m^2)$, coïncidant avec la formule de Balmer et les extensions par Rydberg.

Notons enfin qu'on peut, à partir de ces différentes expression, réécrire la vitesse de l'électron sous la forme:

$$v = c\frac{\alpha}{n} \ . \tag{1.15}$$

1.3 Au delà du modèle de Bohr

Le modèle de Bohr explique donc parfaitement les fréquences des raies de l'hydrogène en dépit du caractère un peu artificiel des hypothèses de départ (orbite circulaire et condition de quantification ad hoc). Il est cependant insuffisant pour expliquer les structures fines observées très tôt dans le spectre. Un pas important est franchi indépendamment par Sommerfeld et Wilson en 1915. Ils considèrent un mouvement elliptique plus général et imposent des conditions de quantification à tous les couples de variables conjuguées ayant une influence sur la dynamique. En termes modernes, ils quantifient le mouvement radial et la norme du moment angulaire. La formule obtenue est en parfait accord avec les structures fines mesurées (et avec la théorie quantique moderne au même ordre d'approximation).

Bien sûr, Bohr, Sommerfeld et bien d'autres cherchent ensuite à adapter cette "première théorie des quanta" à des systèmes atomiques plus complexes. Ils s'intéressent à l'hélium, le plus simple des systèmes complexes avec ses deux électrons (le spectre expérimental est alors bien connu). Et c'est là où le bât blesse! Le système à trois corps ne peut en effet être traité explicitement en mécanique classique. Les règles de quantification utilisées pour l'hydrogène ne sont pas directement applicables. Commence alors un superbe travail de mécanique céleste visant à mettre la version classique du problème sous une forme propre à la quantification. En dépit d'efforts énormes, tirant parti des techniques les plus sophistiquées de la mécanique analytique, toutes ces tentatives échoueront et il ne sera pas possible de donner une interprétation convaincante du spectre de l'hélium. Un grave crise de la mécanique quantique s'ensuivit. Elle ne sera réglée qu'en 1925-1926 par l'invention simultanée de la mécanique des matrices par Heisenberg et de la mécanique ondulatoire par Schrödinger (on reconnaîtra très rapidement l'équivalence des deux approches, en dépit d'un débat plutôt vif initialement entre Heisenberg et Schrödinger).

Nous ne rappellerons pas ici les développements suivants qui ont permis, en particulier avec Dirac, qui introduit la relativité dans le problème, de donner une théorie complète et rigoureuse de l'atome d'hydrogène et, au prix de techniques de calculs complexes, des éléments plus lourds. Notons toutefois

que l'atome d'hydrogène a continué longtemps à constituer une pierre de touche de la mécanique quantique. La mesure et l'interprétation du déplacement du niveau 2S par rapport au $2P_{1/2}$, nul dans le modèle relativiste de Dirac, le fameux "Lamb shift", a joué un rôle essentiel dans le développement de la théorie quantique des champs et des techniques de traitement des infinis.

Pour mémoire, nous rappellerons brièvement ici les résultats de la mécanique quantique standard non relativiste. On en trouvera une dérivation détaillée dans le Cohen. Les fonctions d'onde propres du Hamiltonien quantique (constitué du seul potentiel Coulombien en $1/r$) sont repérées par trois "nombres quantiques" n, ℓ et m. Elles se mettent, en coordonnées sphériques, sous la forme du produit d'une fonction de r par une fonction de la direction angulaire, une "harmonique sphérique": $\Psi_{n\ell m} = R_{n\ell}(r)Y_l^m(\theta,\phi)$. ℓ décrit la longueur du moment cinétique: le carré du moment cinétique vaut $\ell(\ell+1)\hbar^2$. ℓ est un nombre entier positif ou nul. m décrit la projection du moment cinétique sur un axe "de quantification", Oz en l'occurrence. Cette projection vaut simplement $m\hbar$ et est comprise entre $-\ell$ et ℓ. Enfin, n, nombre quantique principal, décrit le nombre des "extrema" (nombre de zéros plus un) de la fonction d'onde radiale $R_{n\ell}$. On montre que ℓ est au plus égal à $n-1$. L'énergie du niveau n, ℓ, m est, à ce degré d'approximation, R/n^2, la valeur prédite par le modèle de Bohr.

On peut donc classer les niveaux par valeurs croissantes du nombre quantique principal. Le niveau $n = 1$, $\ell = m = 0$, fondamental, est appelé 1S. Généralement, les niveaux de $\ell = 0$ sont notés S, $\ell = 1$ P, $\ell = 2$ D et $\ell = 3$ F. Ces notations, anciennes mais universelles, sont descriptives de l'aspect des raies spectrales: S pour "sharp", P pour "principal", D pour "diffuse", F pour "fundamental". Les raies connectant un niveau S au fondamental, S lui aussi, sont interdites par transition dipolaire électrique. Elles ont donc une durée de vie importante et une très faible largeur spectrale. Elles sont aussi relativement insensibles aux champs électriques ou magnétiques parasites. Les raies d'un niveau P vers le fondamental S sont très autorisées et donc très intenses. Les raies des niveaux D sont très sensibles aux champs électriques parasites dans les décharges et apparaissent larges, diffuses, dans les spectres.

Avec ces notations, les premiers niveaux excités sont 2S et 2P ($m = 0, \pm 1$). Les suivants 3S,3P,3D.... Les fonctions d'onde correspondantes, pour mémoire, sont

$$\Psi_{1S} = \frac{1}{\sqrt{\pi a_0^3}} e^{-r/a_0} \tag{1.16}$$

$$\Psi_{2S} = \frac{1}{\sqrt{8\pi a_0^3}} \left(1 - \frac{r}{2a_0}\right) e^{-r/2a_0} \tag{1.17}$$

$$\Psi_{2P,m=\pm 1} = \mp \frac{1}{8\sqrt{\pi a_0^3}} \left(\frac{r}{a_0}\right) e^{-r/2a_0} \sin\theta e^{\pm im\phi} \tag{1.18}$$

$$\Psi_{2P,m=0} = \frac{1}{4\sqrt{2\pi a_0^3}} \left(\frac{r}{a_0}\right) e^{-r/2a_0} \cos\theta \ . \tag{1.19}$$

Ce modèle n'est qu'approché. Il faut lui ajouter les effets relativistes (qui impliquent en particulier l'existence du spin de l'électron) et les effets d'entraînement du noyau décrivant, ensemble, la structure fine. Il faut aussi tenir compte de l'interaction entre l'électron et le spin nucléaire, des effets de volume du noyau (un proton n'est pas ponctuel) et enfin des corrections radiatives qui décrivent le Lamb shift. Tout cela peut être calculé avec une précision remarquable, puisque l'accord théorie/expérience sur le spectre de l'hydrogène atteint maintenant presque les 10^{-12} en valeur relative, précision limitée seulement par la connaissance de la structure du proton, qui joue un rôle essentiel à ce degré d'exactitude, et par la stabilité des horloges étalon utilisées pour la détermination des fréquences. Avec de tels accords, la constante de Rydberg est sans aucun doute la constante la mieux connue de la physique fondamentale.

Ces fonctions d'onde sont bien éloignées du modèle d'orbite circulaire de Bohr. En fait, les niveaux S qui correspondent à un moment cinétique orbital nul seraient représentés en mécanique classique

par une orbite complètement dégénérée en une droite où l'électron rebondit à chaque période sur le noyau. L'incertitude quantique rend, qualitativement, "incertaine" la direction de cette orbite et les fonctions d'onde quantiques des états S sont à symétrie sphérique. Si on peut retrouver des orbites de Bohr, ou quelque chose qui y ressemble un peu, c'est en allant vers les grand nombres quantiques. Le principe de correspondance nous enseigne en effet que les résultats de la mécanique quantique doivent rejoindre ceux de la mécanique classique quand *tous* les nombres quantiques sont grands. Un grand nombre quantique principal (n de l'ordre de quelques dizaines) correspond à ce qu'on appelle maintenant un état de Rydberg. En effet, pour ces niveaux très excités, l'électron orbite très loin du noyau et des autres électrons et tout se passe comme si l'atome était hydrogénoïde, avec des niveaux en $1/n^2$. Les autres nombres quantiques doivent être grands eux aussi: $\ell = |m| = n - 1$. L'orbitale de ces niveaux, dits "circulaires" dans la littérature moderne, est un tore très mince centré sur un cercle de rayon $a_0 n^2$. C'est évidemment l'orbitale la plus proche du modèle de Bohr. Bien sûr, l'électron ne peut être localisé précisément sur cette orbite circulaire et la densité de présence est uniforme le long du périmètre. Cependant, ces niveaux ressemblent beaucoup au modèle de Bohr et de nombreuses prédictions classiques de ce modèle donnent des ordres de grandeur corrects pour les atomes circulaires.

Partie II

Relativité restreinte

Introduction

Nous décrirons dans ce chapitre la théorie de la relativité restreinte, essentiellement telle qu'elle fut formulée par Einstein. Initialement, l'objet de la relativité était de réconcilier l'électromagnétisme et la cinématique : comme nous avons vu dans l'introduction à ce cours, la relativité Galiléenne n'est pas directement compatible avec l'électromagnétisme tel qu'il fut formulé par Maxwell. C'est essentiellement sur l'application à l'électromagnétisme de la relativité restreinte que portera ce chapitre. En revanche, il faut bien voir que la relativité restreinte s'applique dans un beaucoup plus large domaine. Elle joue en particulier un rôle central pour la physique des particules et la physique des accélérateurs. Elle est aussi essentielle en astronomie, beaucoup de sources de rayonnement cosmiques impliquant des déplacements à des vitesse proches de celle de la lumière.

Ce chapitre comportera deux étapes essentielles. Après un bref rappel de la relativité galiléenne, nous constaterons les difficultés que pose l'immersion de l'électromagnétisme dans cette relativité et donc dans la cinématique classique. Nous montrerons en particulier l'incompatibilité grave de l'électromagnétisme avec la loi ordinaire de composition des vitesses. Nous postulerons donc un nouveau principe de relativité, imposant à toutes les lois de la physique, y compris l'électromagnétisme, d'être invariantes dans un changement de référentiel galiléen. La vitesse de la lumière devenant indépendante du référentiel, la loi de composition des vitesses et l'ensemble de la cinématique sont condamnées. Il nous faudra donc d'abord détruire la cinématique et la dynamique[3] newtoniennes telles que nous les connaissons maintenant.

Il nous faudra formuler une nouvelle transformation des coordonnées et du temps décrivant les changements de référentiels, la transformation de Lorentz. Nous verrons en effet, par quelques arguments très simples, qu'un des postulats de base de la mécanique classique, l'universalité du temps et de la simultanéité, doit être abandonné. On mesure peut être assez mal aujourd'hui à quel point la démarche d'Einstein fut audacieuse, remettant en cause les postulats les plus intuitifs de la mécanique. La phase conceptuellement la plus difficile de notre travail, qui fera l'objet du premier chapitre, sera alors terminée.

Le deuxième chapitre, beaucoup plus mathématique que physique, sera consacré à l'introduction de notations tensorielles, bien adaptées à l'espace–temps à quatre dimensions de la relativité. Nous introduirons en particulier des conventions de notations très puissantes, dues à Einstein, qui permettent d'écrire de manière compacte et fiable les expressions parfois complexes auxquelles conduisent les calculs relativistes. Ces notations s'avèrent indispensables pour aborder la relativité générale, théorie géométrique de la gravitation.

Nous formulerons, au chapitre suivant, les lois de la nouvelle dynamique. Nous écrirons en particulier, dans une approche lagrangienne, le lagrangien d'une particule libre et nous en déduirons l'expression de la quantité de mouvement relativiste. Nous démontrerons en passant la formule la plus célèbre de l'histoire de la physique (nous laissons au lecteur le soin de deviner laquelle). Nous donnerons également la forme relativiste du principe fondamental de la dynamique que nous ne pourrons guère exploiter sans une forme explicite des forces, au moins de la force de Lorentz. Nous n'explorerons donc pas très en détails cette partie de la relativité qui se conclura par une brève description de la

[3] Rappelons que la cinématique décrit les mouvement indépendamment de leurs causes et que la dynamique permet de prévoir le mouvement si on en connaît les causes.

théorie relativiste des collisions, d'une grande importance en physique des particules.

Pour un cours centré sur l'électromagnétisme, nous consacrerons en effet l'essentiel de nos efforts au dernier chapitre de cette partie. Nous chercherons à y construire une théorie non triviale d'interaction entre particules transmise par un champ. Nous postulerons des formes simples pour le lagrangien d'interaction et pour le lagrangien décrivant ce champ et nous écrirons les équations de Lagrange correspondantes. Nous constaterons sans déplaisir que la structure de cette théorie de champ est celle de l'électromagnétisme. Nous aurons donc montré à quel point l'électromagnétisme de Maxwell s'adapte naturellement au cadre relativiste. Nous en profiterons pour examiner quelques problèmes simples d'électromagnétisme, du mouvement de particules relativistes dans des champs imposés aux bilans d'énergie–impulsion pour le champ lui même. Nous montrerons ainsi que cette approche complètement relativiste, outre son élégance, permet de dériver des lois importantes qui ne sont accessibles qu'au prix de calculs lourds en électromagnétisme "classique".

Chapitre 1

Cinématique relativiste

La première étape est donc de comprendre les incompatibilités entre électromagnétisme et cinématique classique, et de refonder une cinématique tout à fait nouvelle. Nous allons commencer par quelques très brefs rappels de cinématique galiléenne ou newtonienne.

1.1 Rappels de relativité galiléenne

1.1.1 Transformation de Galilée

Il est très intuitif que le mouvement d'un point dépende de l'observateur. Pour utiliser un vocabulaire ferroviaire[1], le passager de train a une vitesse faible ou nulle par rapport à celle du contrôleur, alors qu'il a une vitesse élevée par rapport au garde barrière.

La notion centrale de la cinématique (classique ou relativiste) est celle de *référentiel*. Un référentiel, c'est un ensemble d'observateurs, immobiles les uns par rapport aux autres. Ces observateurs peuvent constater le passage du mobile à leur position. La connaissance de la position des observateurs concernés permet alors de déterminer la trajectoire du mobile. On peut bien sûr convenir d'un repère (cartésien, orthonormal) pour repérer ces positions au moyen de trois coordonnées. Les observateurs sont de plus munis d'horloges qui leur permettent de noter l'instant auquel le mobile passe en face d'eux, le mouvement étant alors complètement déterminé par la trajectoire et la loi horaire. Ces horloges peuvent être constituées de n'importe quel phénomène physique périodique, suffisamment rapide à l'échelle du mouvement pour en donner une description temporelle convenable. Nous supposerons que toutes les horloges de tous les observateurs d'un même référentiel sont synchronisées (indiquent la même valeur au même instant). Cette synchronisation ne pose aucune difficulté en cinématique classique, puisque temps et espace sont complètement découplés. Il suffit, par exemple, que tous les observateurs se retrouvent en un même point pour faire le zéro de leurs horloges à un moment commun. Certes, ces précautions pour la définition du temps paraissent superfétatoires en cinématique classique. Nous verrons, en revanche, qu'elles sont très importantes en cinématique relativiste.

Un mouvement dans un référentiel $\mathcal{R}$ est alors défini par les trois fonctions $x(t), y(t), z(t)$ représentant la position en fonction du temps commun des observateurs. Le même mouvement serait décrit dans un autre référentiel $\mathcal{R}'$, en mouvement par rapport à $\mathcal{R}$, par trois autres fonctions du temps commun des observateurs de $\mathcal{R}'$: $x'(t'), y'(t'), z'(t')$. En mécanique classique, on admet sans restrictions l'identité des temps (à une synchronisation près) des observateurs de $\mathcal{R}$ et de $\mathcal{R}'$[2]. Il

[1] Les papiers originaux sur la relativité emploient souvent des expériences de pensée utilisant des trains et des gares, parfois même des tunnels. C'est sans doute lié au succès grandissant des transports ferroviaires au début du siècle et à leur importance sociologique. Pour céder à la tradition, nous emploierons ce genre de vocabulaire dans ce cours, bien que les effets relativistes soient complètement négligeables, même avec les trains les plus modernes.

[2] Cette hypothèse était déjà faite explicitement par Newton dans ses *Principia*. S'il en avait tout à fait reconnu l'importance, il n'avait guère de doutes sur sa validité.

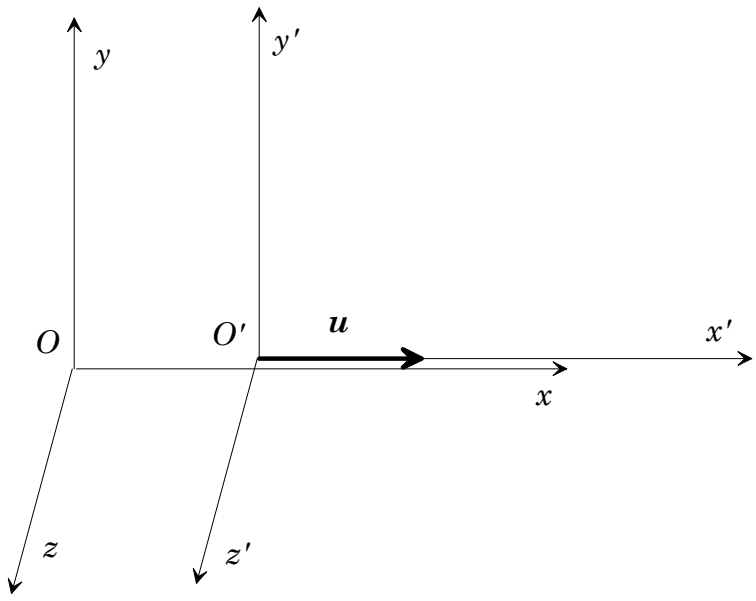

Figure 1.1: Choix des axes dans deux référentiels $\mathcal{R}$ et $\mathcal{R}'$ en mouvement relatif. Les axes des deux repères sont parallèles. Les axes Ox et $O'x'$, alignés avec la vitesse relative $\mathbf{u}$, coïncident à chaque instant.

est possible alors de donner la transformation qui fait se correspondre les mouvement vus dans deux référentiels différents.

Dans le cas le plus simple, où les deux référentiels sont en translation uniforme l'un par rapport à l'autre, cette transformation est la transformation dite de Galilée. Sans restreindre du tout la généralité, on peut choisir les axes dans $\mathcal{R}$ et $\mathcal{R}'$ de telle manière que:

- Les axes Ox et $O'x'$ coïncident a tout instant et sont parallèles à la vitesse $\mathbf{u}$ de $\mathcal{R}'$ par rapport à $\mathcal{R}$.
- Les origines O et O' sont confondues à l'instant $t = 0$.
- Les axes Oy et $O'y'$, d'une part, et les axes Oz et $O'z'$, d'autre part, sont constamment parallèles et coïncident à $t = 0$.

La figure 1.1 présente la géométrie choisie. Nous l'exposons en détail parce que nous choisirons la même pour décrire les changements de référentiel en relativité restreinte.

La loi de transformation de Galilée s'écrit alors trivialement:

$$\left.\begin{array}{l} x'(t) = x(t) - ut \\ y'(t) = y(t) \\ z'(t) = z(t) \end{array}\right\} \tag{1.1}$$

C'est cette transformation, tellement triviale qu'elle est bien rarement écrite explicitement, qui sera remplacée par la transformation de Lorentz en relativité einsteinienne. Cette transformation de Galilée contient, par simple dérivation par rapport au temps, la loi de composition des vitesses:

$$\mathbf{v} = \mathbf{v}' + \mathbf{u} \tag{1.2}$$

(vitesse absolue égale vitesse relative plus vitesse d'entraînement).

La dynamique newtonienne résulte alors du principe d'inertie de Galilée: il existe une classe de référentiels privilégiés, les référentiels galiléens, en mouvement de translation uniforme les uns par rapport aux autres, tels que le mouvement d'une particule libre y soit rectiligne et uniforme.

1.1.2 Les difficultés de la cinématique classique

La loi de composition des vitesses, telle que nous venons de la rappeler, est difficilement compatible avec l'électromagnétisme de Maxwell. La conséquence la plus importante et la plus nouvelle des équations de Maxwell est en effet la prédiction de l'existence d'ondes se propageant à la vitesse c. Le problème qui apparaît immédiatement est celui du référentiel dans lequel cette vitesse est définie, le seul donc dans lequel les équations de Maxwell seraient directement applicables.

Le sentiment le plus naturel, qui prédominait très largement à la fin du siècle dernier, était que les ondes électromagnétiques se propageaient dans un milieu baignant l'univers entier: *l'éther*. L'analogie entre ondes électromagnétiques et ondes sonores était en effet présente à tous les esprits. Les difficultés apparaissent toutefois très vite dès qu'on examine les propriétés de cet hypothétique éther. Il doit en effet être omniprésent et infiniment rigide pour propager des ébranlement transverses à grande vitesse. Mais il doit, dans le même temps, être impondérable et infiniment perméable au mouvement des corps matériels (puisque, par exemple, l'étude sur quelques siècles de la rotation terrestre ne révèle aucun frottement). Ce "fluide" si particulier se trouvait ainsi doté de propriétés presque aussi extraordinaires que le calorique du siècle précédent ou, encore avant lui, le phlogistique.

Il existe aussi une difficulté philosophique grave avec l'introduction de l'éther. Les physiciens avaient mis plus de 20 siècles, entre Aristote et Copernic, pour comprendre que notre petite planète n'est pas le centre de l'univers. Le principe de relativité selon Galilée avait le mérite d'indiquer qu'aucun référentiel galiléen n'est particulièrement privilégié. L'introduction de l'éther devait briser cette "démocratie" des référentiels en introduisant un référentiel très particulier, celui de l'éther, le seul dans lequel les équations de Maxwell devaient s'appliquer. Ce genre d'argument a certainement joué un rôle essentiel pour Einstein.

Les dernières difficultés, les plus graves en pratique, mais qui n'ont pas forcément joué le rôle majeur qu'on leur attribue généralement dans la genèse de la relativité, sont d'ordre expérimental. Si la vitesse de la lumière est définie dans le référentiel de l'éther et si elle obéit à la loi de composition des vitesses, on doit pouvoir mesurer une variation de cette vitesse pour des mouvements assez rapides par rapport à l'éther. Le mouvement de la terre sur son orbite autour du soleil est suffisamment rapide (30 km/s) pour que la variation soit mesurable dans une expérience d'interférométrie optique sensible. La célèbre expérience de Michelson fut conçue dans ce but. D'une sensibilité tout à fait remarquable pour l'époque, encore honorable aujourd'hui, elle aurait dû mettre clairement en évidence le mouvement de la terre par rapport à l'éther[3]. Or cette expérience fut tout à fait négative (ou plutôt très positive): la vitesse de la lumière semblait indépendante du mouvement de la terre par rapport au soleil.

On pouvait, devant ce résultat négatif, adopter deux points de vue. Le premier était de tenter de "réparer" la théorie de l'éther. Si on ne pouvait décemment supposer que le référentiel absolu était celui de la terre (la révolution copernicienne était passée par là), on pouvait supposer que l'éther était entraîné au voisinage des corps massifs, une analogie évidente avec l'entraînement de la couche limite en hydrodynamique. On pouvait aussi supposer, avec Lorentz, une "contraction" de la longueur des objets matériels dans la direction du mouvement, fondée sur une théorie électrostatique des interactions entre particules dans la matière. On pouvait supposer aussi un lien entre la vitesse de la lumière et celle de sa source (les sources utilisées par Michelson étant liées à son appareil). Si de telles modifications "ad hoc" de l'électromagnétisme permettaient d'expliquer le résultat négatif de l'expérience de Michelson, ils ne constituaient pas un corps théorique cohérent. Il était à craindre que de nouvelles modifications tout aussi arbitraires ne doivent être apportées au gré des résultats expérimentaux et que l'électrodynamique ne finisse, comme la théorie astronomique des cycloïdes, en un corps raffiné de règles arbitraires qui décrivent correctement mais ne prédisent rien.

L'autre attitude, beaucoup plus courageuse puisqu'elle conduit, comme nous le verrons, à mettre en cause des notions très fondamentales, était d'admettre que la vitesse de la lumière n'obéissait pas à

[3]Nous ne détaillerons pas ici le principe de cette expérience: cette description n'est pas indispensable pour la suite de l'exposé. Le lecteur intéressé pourra trouver une description détaillée dans pratiquement tous les manuels de relativité.

la loi de composition des vitesses. Cela impliquait bien sûr que la cinématique galiléenne était erronée (ou, du moins, n'était qu'une approximation valide pour des vitesses petites devant celle de la lumière) et donc que toute la physique était à reconstruire (sauf, peut être, l'électrodynamique). C'est la voie que suivit Einstein avec le succès que l'on connaît et qu'il ouvrit par son célèbre article de 1905: "Sur l'électrodynamique des corps en mouvement"[4]. Le principe fondamental de cette nouvelle physique, le "principe de relativité" est exposé dans le prochain paragraphe.

1.2 Principe de relativité

1.2.1 Enoncé

Il existe une classe de référentiels privilégiés, en translation uniforme les uns par rapport aux autres (que nous continuerons à appeler "référentiels galiléens"), dans lesquels toutes les lois de la physique prennent la même forme.

Si toutes les lois de la physique prennent la même forme, les équations de Maxwell sont valides dans tous les référentiels et la vitesse de la lumière, c, est la même dans tous les référentiels. Nous centrerons cet exposé sur l'invariance de la vitesse de la lumière. Il serait tout aussi possible de ne pas faire jouer un rôle aussi central à l'électromagnétisme. On pourrait simplement postuler qu'il existe une vitesse limite de propagation de toutes les interactions et effectuer tous les raisonnements qui vont suivre sur cette vitesse. Il suffirait, enfin, de constater qu'expérimentalement la vitesse de la lumière dans le vide est identique à la vitesse limite, à la précision des mesures. Il n'est pas absolument exclus, en effet, bien que cela soit très peu vraisemblable, que le photon possède une très petite masse, rendant la vitesse de la lumière très légèrement inférieure à la vitesse limite qui apparaît en relativité.

Ce principe de relativité, de prime abord, semble ne rien remettre en cause d'essentiel et semble très voisin du principe de relativité de la physique classique. Il n'en est rien, comme nous allons le voir en considérant deux expériences de pensée. Nous allons montrer en effet que le principe de relativité a deux conséquences immédiates:

- Le temps ne s'écoule pas de la même façon dans deux référentiels galiléens en mouvement relatif (deux horloges en mouvement relatif bâties sur le même modèle ne battent pas au même rythme).
- Deux événements qui se produisent simultanément dans un référentiel peuvent se produire à des instants différents dans un autre référentiel.

Remettre en cause des propriétés aussi intuitives de l'espace et du temps ne sera pas sans conséquences. Il est clair, en particulier, que la transformation appelée à remplacer la transformation de Galilée devra renoncer au caractère absolu du temps et mélanger les coordonnées spatiales et temporelles.

1.2.2 Deux expériences de pensée

Nous considérons donc deux référentiels en mouvement relatif, avec la géométrie décrite dans la figure 1.1. Le référentiel $\mathcal{R}'$ sera celui du contrôleur, ou du train, pour reprendre nos analogies ferroviaires, le référentiel $\mathcal{R}$ celui du chef de gare. Le contrôleur, situé en O', envoie à $t' = 0$ (nous ne confondrons pas les temps dans les deux référentiels) une impulsion lumineuse de durée négligeable dans la direction y' vers un miroir situé en $y' = L$ (voir figure 1.2). L'impulsion, réfléchie par le miroir, revient vers le contrôleur et l'atteint au bout d'un temps $T' = 2L/c$ (nous supposerons, pour ce paragraphe seulement, que, pour la cinématique classique, la vitesse de l'impulsion est c dans $\mathcal{R}'$ - l'ensemble

[4]Nous ne saurions trop recommander la lecture de cet article, ainsi que celle d'un article de revue rédigé dès 1907, qui constitue un exposé très pédagogique de la relativité (Edition de œuvres essentielles d'Einstein, Relativités I—éditions Seuil–CNRS).

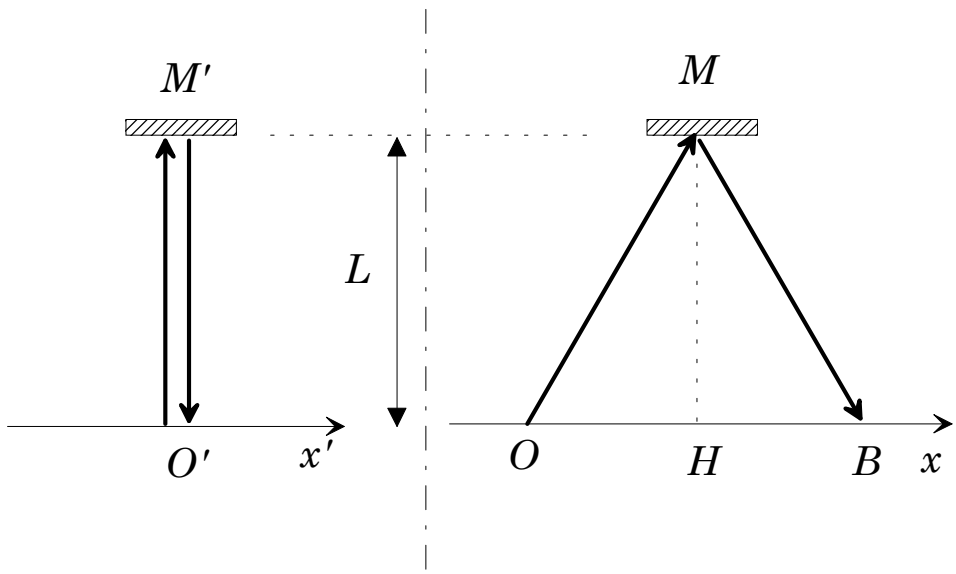

Figure 1.2: Une expérience de pensée établissant le caractère relatif du temps. Un signal lumineux est émis depuis l'observateur O', le long de l'axe $O'z'$, vers un miroir M. Après réflexion sur ce miroir, le signal revient à l'observateur O'. La même expérience est vue, à gauche, dans le référentiel du train et, à droite, dans un référentiel immobile. Pendant l'expérience, l'observateur O' est passé de O à B.

de l'argument pouvant être transposé sans difficultés au cas, plus naturel, où on supposerait que la vitesse de l'impulsion est c dans $\mathcal{R}$). Notons que le contrôleur pourrait ainsi construire une horloge. Renvoyant une deuxième impulsion à l'instant précis où il reçoit la première, il établirait un phénomène périodique et donc une horloge.

Regardons maintenant cette même expérience avec l'oeil du chef de gare (partie droite de la figure 1.2). A $t' = 0$, le contrôleur est en O' et donc aussi en O. De son côté, le miroir s'est déplacé avant que l'impulsion ne l'atteigne. Il occupe donc une position M, à une certaine distance de O sur l'axe Ox. Enfin, le train continue à se déplacer pendant le retour de l'impulsion et le contrôleur occupe la position B au moment du retour. La trajectoire de l'impulsion dans $\mathcal{R}$ est triangulaire.

Imaginons d'abord que le chef de gare ait été nommé avant 1905 et soit donc un adepte de la cinématique classique. Pour lui, la vitesse de la lumière obéit à la loi de composition. La vitesse de l'impulsion a donc une composante $\pm c$ sur Oy et une composante u sur Ox (u étant la vitesse du train). Son module est donc $\sqrt{c^2 + u^2}$. La durée du parcours OM étant L/c (on admettra dans toute la suite que la position du miroir selon y n'est pas affectée par le changement de référentiel; nous en donnerons plus tard une justification détaillée), on a $OH = uL/c$ et la longueur OM vaut $L\sqrt{c^2 + u^2}/c$. Le temps du parcours OM est donc L/c et la durée totale de l'expérience $T = 2L/c$ est identique à celle vue par le contrôleur (la distance parcourue dans $\mathcal{R}$ est plus grande, mais le module de la vitesse est augmenté dans la même proportion). On retrouve bien, naturellement, le postulat d'universalité du temps.

Considérons maintenant un chef de gare ayant admis la validité du postulat de relativité. Pour lui, la vitesse de l'impulsion est c. Le temps de parcours OM est donc OM/c. Il en déduit $OH = uOM/c$. Comme $OM^2 = L^2 + OH^2$, on a $OM = L/\sqrt{1 - u^2/c^2}$. Il obtient donc finalement:

$$T = \gamma T' \ , \tag{1.3}$$

avec

$$\gamma = \frac{1}{\sqrt{1 - \dfrac{u^2}{c^2}}} \ . \tag{1.4}$$

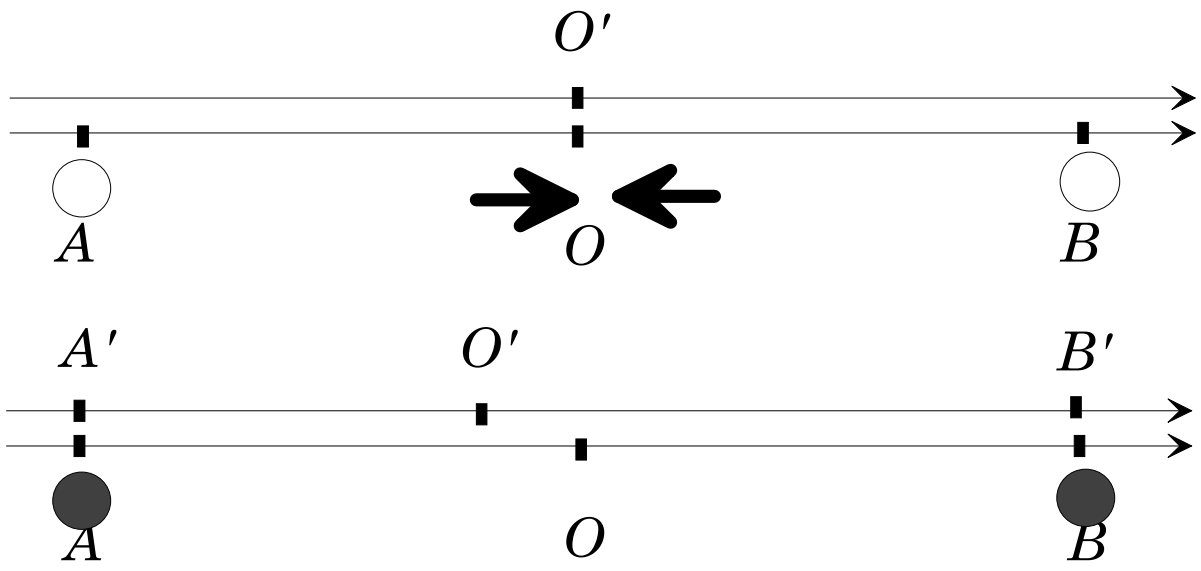

Figure 1.3: Deuxième expérience de pensée illustrant le postulat de relativité. En haut, vue de la situation au moment où le chef de gare, situé en O, et le contrôleur, en O' voient arriver simultanément les signaux lumineux émis par A et B. En bas, situation au moment où les signaux se sont allumés. O' n'est pas encore arrivé en O. Les signaux s'allument en face des observateurs A' et B'.

Le facteur γ (que nous aurons de nombreuses occasions de retrouver) est toujours supérieur à un. La durée de l'expérience mesurée par le chef de gare est plus longue que celle mesurée par le contrôleur. Si chacun construisait une horloge avec le même dispositif, celle du chef de gare battrait plus vite et avancerait par rapport à celle du contrôleur (un cauchemar pour le respect des horaires; heureusement, l'effet est petit comme on pourra le vérifier). Le postulat de relativité a pour conséquence immédiate que le temps n'est pas une notion universelle.

Cette expérience de pensée nous fournit une autre indication sur ce que sera la cinématique relativiste. Le facteur γ n'existe que si la vitesse relative des deux référentiels est plus petite que c. Si ce n'était pas le cas, l'impulsion lumineuse qui se réfléchit normalement sur le miroir dans $\mathcal{R}'$ n'arriverait jamais à rattraper ce miroir dans $\mathcal{R}$, puisque sa vitesse n'est "que" c. Un événement (la réflexion) se produirait dans un référentiel et pas dans un autre, ce qui est bien sûr absurde. Deux référentiels galiléens ne peuvent donc être animés l'un par rapport à l'autre d'une vitesse supérieure (ou même égale) à c. Comme on peut attacher un référentiel galiléen à tout objet en translation uniforme, c apparaît aussi comme une vitesse limite pour tous les objets matériels.

La seconde expérience de pensée que nous allons étudier nous emmènera encore plus loin, puisqu'elle nous forcera à renoncer au caractère absolu de la simultanéité. Nous utiliserons encore les services du chef de gare et du contrôleur embarqué sur son train. Le chef de gare est situé en O, à mi–chemin de deux signaux lumineux A et B. A l'instant $t = 0$, il voit ces deux signaux s'allumer simultanément. S'il sait, ou s'il mesure, que la même distance L le sépare des deux signaux, il en déduira qu'ils se sont allumés simultanément à l'instant $t = -L/c$.

Au même instant $t = t' = 0$, le contrôleur, situé en O', passe devant le chef de gare. Il voit donc, à cet instant précis, les deux signaux A et B s'allumer. Comment peut-il en déduire l'instant auquel ils se sont allumés (dans son échelle de temps, bien sûr)? Il lui faut d'abord déterminer *où* les deux signaux se sont allumés dans son référentiel. Pour cela, il peut parcourir son train et rechercher les deux voyageurs A' et B' (les observateurs) qui étaient juste en face des signaux quand ils se sont allumés. Il pourra ensuite leur demander à quel instant cet événement s'est produit ou utiliser leur position et la vitesse de la lumière pour calculer cet instant.

Supposons d'abord que la cinématique classique s'applique. Le temps étant universel, les deux signaux s'allument dans $\mathcal{R}'$ au même instant $t' = -L/c$. A cet instant O' est à une abscisse $-uL/c$ par rapport à O. Les passagers A' et B' sont alors situés respectivement en $x' = -L + uL/c$ et $x' = L + uL/c$ (voir la partie inférieure de la figure 1.3). La distance $A'O' = L(c-u)/c$ est donc inférieure à la distance $B'O' = L(c+u)/c$. En revanche, la vitesse de l'impulsion venant de A' est $c-u$ et la vitesse de l'impulsion venant de B' est $c+u$. En appliquant la procédure décrite ci–dessus, le contrôleur établira donc que les deux signaux se sont allumés au même instant. L'un était plus proche, mais la vitesse de l'autre impulsion était plus grande.

Que se passe-t-il maintenant si nous appliquons le principe de relativité? La vitesse des deux impulsions lumineuses est la même. Nous ne pourrons pas, pour le moment, déterminer la position des observateurs A' et B' (il nous faudra la transformation de Lorentz pour cela). Nous pouvons comprendre, en revanche, que la distance $A'O'$ est nécessairement inférieure à la distance $O'B'$. Le temps de parcours de l'impulsion provenant de A est donc inférieur à celui de l'autre. Le contrôleur en déduira que le signal A s'est allumé *après* le signal B. Deux événements peuvent être vus comme simultanés ou non par des observateurs appartenant à des référentiels différents. Nous verrons bientôt qu'heureusement cet abandon de l'universalité de la simultanéité ne compromet pas la causalité.

Nous avons vu émerger, en discutant ces deux expériences de pensée, deux des notions essentielles de la relativité: l'événement et l'intervalle.

1.3 Evénements et intervalles

1.3.1 Evénements

Comme nous venons de le voir, le temps n'est plus universel et n'est plus séparable des cordonnées spatiales. Il faudra décrire les expériences en termes d'événements (très littéralement: il s'est passé quelque chose quelque part). Un événement, c'est par exemple l'allumage du signal A ou la réflexion de la lumière sur le miroir dans notre première expérience de pensée. Un événement existe indépendamment du choix du référentiel. On peut caractériser un événement, dans un référentiel donné, par l'observateur qui était sur place (le passager A') et par l'instant, mesuré sur l'horloge de cet observateur, où l'événement s'est produit[5]. On pourra donc complètement caractériser l'événement par quatre nombres: les trois coordonnées spatiales de l'observateur (on se munit d'un repère convenable) et le temps. On décrira donc un événement par la donnée d'un référentiel et d'un quadruplet de nombres (ct, x, y, z) (nous développerons au chapitre suivant des notations tensorielles puissantes pour traiter ces quadruplets). Bien sûr, les coordonnées spatio–temporelles du même événement dans un autre référentiel sont différentes et l'essentiel de notre tâche sera de donner la loi de transformation qui remplace et étend la transformation de Galilée. Il y a un parallèle très fort entre la différence entre événement (indépendant du référentiel) et coordonnées spatio–temporelles et celle qui existe entre un vecteur (indépendant du repère) et ses composantes sur une base donnée.

Nous utiliserons souvent des représentations géométriques des événements. On peut en effet les représenter comme un point dans un espace à quatre dimensions. Cette représentation posant quelques problèmes techniques, on se cantonne souvent à une dimension d'espace. On représente alors un événement comme sur la figure 1.4. Pour des raisons de commodité, on porte sur l'axe vertical le produit ct. Les deux coordonnées dans cet espace ont ainsi la même dimension[6].

[5] On supposera encore que tous les observateurs d'un même référentiel peuvent synchroniser leurs horloges. Il leur est interdit de se déplacer, mais on peut procéder de façon plus subtile. On peut, par exemple, déterminer par des moyens géométriques le milieu du segment AB joignant deux observateurs. On peut placer en ce point une source lumineuse qui s'allume à un certain instant. Si les observateurs A et B font le zéro de leurs horloges au moment où ils voient cette source s'allumer, ils auront établi leur synchronisme.

[6] Les professionnels de la relativité prennent souvent $c = 1$, ce qui simplifie énormément les écritures. A notre niveau, il est peut être imprudent de se priver d'un moyen de vérifier l'homogénéité de nos formules

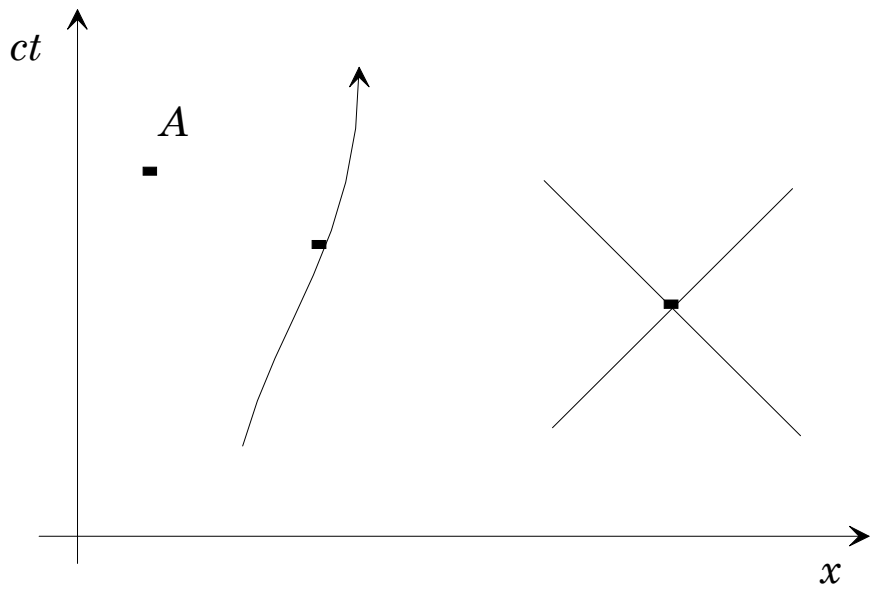

Figure 1.4: Un événement, une ligne d'univers et un cône de lumière. Un événement est représenté par un point dans un espace x, ct. Une ligne d'univers est l'ensemble des événements correspondant aux positions successives d'une particule. Le cône de lumière d'un événement est constitué des lignes d'univers d'un signal lumineux passant par cet événement.

On peut considérer le mouvement d'un point dans un référentiel comme une suite d'événements (la suite des observateurs devant lesquels la particule est passée associée aux instants correspondants). Une telle suite continue d'événements forme dans l'espace-temps une ligne, que nous nommerons "ligne d'univers" de la particule. Une telle ligne est représentée sur la figure 1.4.

La ligne d'univers d'une particule qui se déplacerait à la vitesse de la lumière serait parallèle, dans notre représentation graphique, à la première ou à la deuxième bissectrice. Dans l'espace à quatre dimensions, l'ensemble des lignes d'univers partant d'un point et correspondant à un mouvement à c forme le "cône de lumière" de cet événement (voir aussi la figure 1.4). Les événements antérieurs à l'événement de référence forment le passé du cône de lumière, les autres le futur. Comme c est une vitesse limite, toutes les lignes d'univers passant par un événement donné doivent être à l'intérieur du cône de lumière. Deux événements ne pourront être reliés par un signal ou une relation causale, que s'ils sont dans le cône de lumière l'un de l'autre. Il est évident géométriquement que cette relation est symétrique: si A est dans le cône de lumière de B, alors B est dans le cône de lumière de A. En revanche, cette relation n'est pas transitive dans le cas général, comme on pourra s'en persuader aisément. Si C est dans le passé du cône de lumière de B, lui même dans le futur du cône de lumière de A, alors C n'est pas nécessairement dans le cône de lumière de A. En un mot, si A et C peuvent tous deux être la cause de B, il n'y a aucun lien de causalité a priori entre eux. En revanche, si C est dans le futur de B, il est nécessairement dans le cône de lumière de A: si A est la cause de B qui est lui même la cause de C, alors A peut être la cause de C.

En ces termes, la version relativiste de la causalité apparaît très clairement. Si la physique classique admet qu'un événement puisse être la cause d'un autre s'il lui est antérieur (admettant ainsi implicitement les actions instantanées à distance), la relativité exige que l'un des événements soit effectivement antérieur à l'autre (nous verrons dans le prochain paragraphe que la notion d'antériorité est indépendante du référentiel) mais aussi que les deux événement puissent être reliés par un signal. Nous allons maintenant pouvoir affiner beaucoup ces notions en introduisant l'intervalle.

1.3.2 Intervalle. Invariance de l'intervalle

Considérons deux événements repérés, dans un référentiel donné, par (ct_1, x_1, y_1, z_1) et (ct_2, x_2, y_2, z_2). Si ces deux événements sont sur le cône de lumière l'un de l'autre, ils peuvent être reliés par un signal lumineux se propageant à la vitesse c. On a donc dans ce cas:

$$c^2(t_1 - t_2)^2 = (x_1 - x_2)^2 + (y_1 - y_2)^2 + (z_1 - z_2)^2 \ . \tag{1.5}$$

L'écriture de cette relation suggère d'introduire l'intervalle entre deux événements quelconques par:

$$s_{1,2}^2 = c^2(t_1 - t_2)^2 - (x_1 - x_2)^2 - (y_1 - y_2)^2 - (z_1 - z_2)^2 \ . \tag{1.6}$$

Notons que le choix du signe + pour la composante temporelle de l'intervalle est tout à fait arbitraire. C'est cependant le plus répandu aujourd'hui. L'intervalle jouera le rôle d'une distance dans notre espace–temps à quatre dimensions. Sa seule propriété évidente à ce point est de s'annuler quand les deux événements sont sur le cône de lumière l'un de l'autre. Cette propriété est indépendante du référentiel: le fait pour deux événements d'être ou non reliés par un signal lumineux ne dépend pas de la description du mouvement. Un intervalle nul est donc un invariant dans un changement de référentiel.

Nous établirons rigoureusement, à partir de la transformation de Lorentz, le fait que l'intervalle est indépendant du référentiel (est un "invariant relativiste"), même s'il n'est pas nul. Nous allons donner ici une indication de ce fait par un raisonnement qui, bien qu'il ne soit pas tout à fait rigoureux (il fait appel à des hypothèses supplémentaires implicites), établit de manière simple l'invariance de l'intervalle. Cette invariance nous permettra, dans les prochains paragraphes, de comprendre beaucoup de propriétés de la transformation de Lorentz avant même d'en écrire la forme explicite.

Pour cela, considérons deux événements infiniment voisins. L'intervalle, lui aussi infinitésimal, entre ces événements s'écrit alors, dans un référentiel $\mathcal{R}$:

$$ds^2 = c^2dt^2 - dx^2 - dy^2 - dz^2 \ . \tag{1.7}$$

Considérons les deux mêmes événements dans un autre référentiel $\mathcal{R}'$. L'intervalle entre eux s'écrit

$$ds'^2 = c^2dt'^2 - dx'^2 - dy'^2 - dz'^2 \ . \tag{1.8}$$

On doit pouvoir écrire l'intervalle dans le nouveau référentiel comme une fonction de celui dans $\mathcal{R}$, fonction qui s'annule avec son argument (parce qu'un intervalle nul est conservé). On doit pouvoir développer cette fonction au premier ordre pour les intervalles infinitésimaux que nous manipulons et écrire:

$$ds'^2 = a\, ds^2 \ , \tag{1.9}$$

où a est une constante ne dépendant que de la vitesse relative $\mathbf{u}$ des deux référentiels. En fait l'isotropie de l'espace impose que a ne dépende que du module u de la vitesse $\mathbf{u}$. Considérons maintenant un troisième référentiel $\mathcal{R}''$, en mouvement à la vitesse $\mathbf{v}$ par rapport à $\mathcal{R}$ et $\mathbf{w}$ par rapport à $\mathcal{R}'$. L'intervalle infinitésimal dans ce référentiel, ds''^2, est tel que:

$$ds''^2 = a(v)ds^2 = a(w)ds'^2 = a(w)a(u)ds^2 \ . \tag{1.10}$$

La fonction a doit donc vérifier, pour tout triplet de vitesses relatives:

$$a(w) = \frac{a(v)}{a(u)} \tag{1.11}$$

ce qui est manifestement impossible (le module de la vitesse w dépend de l'orientation relative des deux autres et pas seulement de leur module), à moins que $a = 1$. On établit ainsi l'invariance des intervalles infinitésimaux. Tout intervalle pouvant être obtenu par une intégration d'intervalles infinitésimaux entre les deux événements, on établit ainsi l'invariance d'un intervalle arbitraire.

1.3.3 Conséquences de la conservation de l'intervalle. Causalité relativiste

Si l'intervalle est un invariant relativiste, son signe l'est aussi, bien sûr. Nous distinguerons donc deux types d'intervalles:

- Si $s_{1,2}^2 > 0$, nous dirons que nous avons à faire à un intervalle du genre *temps*. La différence entre les temps des deux événements est supérieure à la distance (en unités convenables). Cela signifie que les deux événements peuvent être reliés par un signal se propageant moins vite que la lumière et qu'il peut donc y avoir un lien de causalité entre eux. En d'autres termes, aussi, les deux événements sont dans le cône de lumière l'un de l'autre.

- Si $s_{1,2}^2 < 0$, la distance spatiale entre les deux événements est plus grande que la distance temporelle. Aucun signal ne peut donc avoir relié les deux événements, ce qui exclut tout lien de causalité (souvenons nous qu'aucune interaction ne peut se propager plus rapidement que la vitesse limite c). Nous dirons alors que nous avons à faire à un intervalle de genre "espace".

- Si $s_{1,2}^2 = 0$, les deux événements peuvent avoir été reliés par un signal se propageant à la vitesse de la lumière. Nous dirons alors que l'intervalle est du genre "lumière".

L'intervalle, ou du moins son signe, est très fortement relié à la notion de causalité. Il est donc essentiel que la nouvelle cinématique prédise l'invariance de l'intervalle, de manière que les liens de causalité entre événements soient indépendants des observateurs. La causalité classique, qui n'exigeait que des relations d'antériorité entre la cause est la conséquence exige maintenant deux conditions. D'abord, la cause et la conséquence doivent être dans le cône de lumière l'une de l'autre pour qu'une interaction ait eu le temps de se propager entre elles (la notion d'interaction instantanée à distance, commune en mécanique classique, disparaît en relativité). D'autre part, il faut encore que la cause précède la conséquence. Il est donc important que les notions de passé et de futur, à l'intérieur du cône de lumière, soient elles aussi des invariants relativistes.

Pour établir cette invariance, considérons le cône de lumière de l'événement O et un événement M dans ce cône de lumière. Nous avons défini le futur de O comme l'ensemble des événements du cône de lumière de O de coordonnée temporelle supérieure à celle de O et nous supposerons M situé dans cette partie du cône de lumière. Si, dans un changement de référentiel, M passait dans le passé de O, cela impliquerait qu'il existe un changement de référentiel pour lequel M et O soient confondus. En effet, quel que soit le changement de référentiel, M reste dans le cône de lumière de O. Par continuité, passer du futur au passé de O impose qu'il existe un changement de référentiel amenant M et O à coïncidence. Mais ceci est contraire à l'invariance de l'intervalle, qui deviendrait nul dans ce changement de référentiel, alors qu'il ne l'est pas initialement. Nous en déduirons donc que les notions de passé et de futur sont des invariants relativistes, ce qui est d'une importance cruciale pour que la causalité garde un sens en relativité. Notons que cette invariance ne tient que pour deux événements situés dans le cône de lumière l'un de l'autre. Si ce n'est pas le cas et si les deux événements ne peuvent être reliés par aucun lien de causalité, l'ordre des temps peut être modifié par un changement de référentiel (c'est par exemple le cas dans l'expérience de pensée du train et des deux signaux que nous avons détaillée plus haut). Bien sûr, nous préciserons quantitativement ces notions dans le paragraphe suivant quand nous disposerons de la forme explicite de la transformation de Lorentz.

Notons enfin, pour finir, que tous les intervalles pris sur la ligne d'univers d'une particule matérielle sont du genre temps.

1.3.4 Temps propre

Nous pouvons appliquer l'invariance de l'intervalle au problème des horloges en mouvement que nous avons déjà abordé dans notre première expérience de pensée. Nous y avons vu que la période d'une horloge (l'aller et retour d'un signal lumineux) n'était pas la même pour le contrôleur et le chef de gare.

Nous allons établir ce résultat de manière plus générale en utilisant l'invariance de l'intervalle. Là encore, toutes ces notions seront beaucoup plus précises quand nous disposerons de la forme explicite de la transformation de Lorentz.

Considérons donc une particule, ou une horloge, en mouvement arbitraire par rapport à un référentiel $\mathcal{R}$. Si ce mouvement est accéléré, il n'existe pas de référentiel galiléen dans lequel la particule soit au repos à tout instant. En revanche, on peut considérer à chaque instant le référentiel galiléen dont la vitesse $\mathbf{v}$ coïncide avec celle de la particule. Nous appellerons ce référentiel le référentiel tangent au mouvement $\mathcal{R}'$. A l'instant considéré, on peut faire en sorte que la particule soit située à l'origine O' de $\mathcal{R}'$, avec une vitesse nulle.

Considérons maintenant un intervalle de temps infinitésimal dt dans $\mathcal{R}$. Pendant ce temps, la particule se déplace de $d\mathbf{l} = \mathbf{v}dt$. Les deux événements correspondant aux deux extrémités de ce mouvement infinitésimal sont donc séparés par un intervalle

$$ds^2 = c^2dt^2 - dl^2 = c^2dt^2\left(1 - \frac{v^2}{c^2}\right) \ . \tag{1.12}$$

Considérons maintenant les deux mêmes événements dans $\mathcal{R}'$, le référentiel tangent. Dans ce référentiel, la vitesse de la particule est nulle. Son déplacement est donc nul au premier ordre en dt'. L'intervalle s'écrit donc aussi:

$$ds^2 = c^2dt'^2 \ . \tag{1.13}$$

Nous appellerons donc "temps propre" l'intervalle de temps s'écoulant dans $\mathcal{R}'$ et nous le noterons $d\tau = dt'$. En rapprochant les deux expressions précédentes de l'intervalle, nous pouvons écrire

$$dt = \gamma d\tau \tag{1.14}$$

avec

$$\gamma = \frac{1}{\sqrt{1 - \frac{v^2}{c^2}}} \ . \tag{1.15}$$

Nous retrouvons, de façon plus générale, que l'intervalle de temps mesuré dans le référentiel tangent ou le référentiel propre dans le cas d'un mouvement uniforme, est plus court que l'intervalle mesuré dans un autre référentiel. Le facteur de "dilatation" du temps, γ, toujours supérieur à un, a l'expression que nous avions déjà trouvée au paragraphe précédent. Il est important de constater à ce point que $d\tau$ est une quantité indépendante de l'observateur. Tout observateur, indépendamment de son état de mouvement par rapport à la particule, peut calculer un intervalle de temps propre à partir d'un intervalle de temps dans son référentiel, et en déduire, par intégration, le temps propre de la particule entre deux événements servant de référence. Le résultat obtenu sera le même pour tous les observateurs (si ils choisissent les mêmes événements de référence). Le temps propre est donc une propriété intrinsèque de la particule. Nous nommerons une telle quantité un 4–scalaire dans le chapitre suivant.

Le raisonnement est fait ici pour des intervalles de temps infinitésimaux. On peut le généraliser à des intervalles arbitraires. A chaque instant, on peut définir un référentiel tangent pour la particule. On peut alors déterminer le "temps propre" de la particule, τ, en intégrant les intervalles de temps propre infinitésimaux. En intégrant aussi la relation entre temps propre et intervalle de temps dans $\mathcal{R}$, on montrera que la durée propre est toujours inférieur à la durée mesurée dans $\mathcal{R}$.

Cette "dilatation des temps" a plusieurs conséquences pratiques mesurables qui ont apporté des confirmations éclatantes au principe de relativité.

Considérons d'abord une particule instable de durée de vie (moyenne) T. Dans quel référentiel doit-on utiliser cette durée de vie? Elle n'a bien sûr de signification que dans le référentiel de la particule. C'est en effet une "horloge" interne à la particule qui déclenche sa désintégration. Dans le

référentiel du laboratoire, $\mathcal{R}$, la durée de vie moyenne de la particule sera alors γT (nous supposerons, pour fixer les idées, que la particule est en mouvement rectiligne uniforme – γ est donc une constante). Si la vitesse de la particule est très proche de celle de la lumière, la facteur de dilatation temporelle γ est très grand devant un et la durée de vie "vue" dans le référentiel du laboratoire est très grande par rapport à la durée de vie intrinsèque. C'est cet effet qui permet d'observer, dans les chambres à bulles ou à fils, les traces de particules à durée de vie très courte.

La dilatation des temps se manifeste aussi à une échelle de vitesses plus accessible, à condition de disposer d'horloges de haute précision. Le réseau d'horloges atomiques qui fixent le temps international doit périodiquement être recalé. Pour cela, on transporte physiquement d'un site à l'autre des horloges portables de haute précision. A l'arrivée du voyage, il faut corriger l'horloge mobile de la dilatation relativiste des temps[7] que l'on pourra estimer numériquement à titre d'exercice.

Citons également le célèbre "paradoxe des jumeaux", dû à Langevin. De deux frère jumeaux, l'un reste sur terre et l'autre vole vers Proxima du centaure, à une distance de 4 années lumière, avec une vitesse constante, proche de celle de la lumière. A peine arrivé, le jumeau voyageur fait demi-tour et revient sur Terre à la même vitesse. A l'arrivée, le temps écoulé pour le jumeau terrestre est de huit ans (4 ans pour l'aller, autant pour le retour). En revanche, pour le voyageur, le temps écoulé n'est que de $8/\gamma$ ans, beaucoup plus court. Nous verrons, quand nous aurons explicité la forme de la transformation de Lorentz, que le jumeau voyageur voit l'étoile de destination s'approcher de lui à une vitesse proche de celle de la lumière. En revanche, elle est initialement beaucoup plus proche de lui que 4 années-lumière. Le jumeau voyageur revient donc sur terre plus jeune que son frère! Le paradoxe apparent est qu'il semble y avoir une parfaite symétrie entre les deux jumeaux, incompatible avec cette différence d'âge: dans le référentiel du voyageur, le jumeau terrestre s'éloigne et se rapproche à grande vitesse.

La "solution" de ce paradoxe apparent est que le référentiel du jumeau voyageur n'est pas un référentiel galiléen. Le temps propre tel que nous l'avons défini n'est pas le temps mesuré dans un référentiel donné. C'est une accumulation de temps infinitésimaux tous calculés dans des référentiels galiléens différents, les référentiels tangents au mouvement accéléré du mobile. Considérer la situation du point de vue du jumeau voyageur reviendrait à définir un temps pour un référentiel bien défini (celui du jumeau voyageur), accéléré. Ceci n'est pas possible dans le cadre de la relativité restreinte. En relativité générale, le temps est affecté par la gravitation ou de manière équivalente par l'accélération. On peut alors effectivement définir un temps pour le jumeau voyageur et retrouver rigoureusement la dissymétrie entre les deux jumeaux.

Cette brève étude du paradoxe des jumeaux introduit naturellement la notion de célérité. Si on désire voyager loin, ce qui importe c'est le temps propre utilisé (celui dans lequel on vieillit) et la distance parcourue dans le référentiel immobile. On peut définir alors une vitesse, que nous nommerons célérité, en termes du temps estimé dans un référentiel et de l'espace estimé dans un autre. De façon évidente, la célérité est le produit de la vitesse ordinaire par le facteur γ de dilatation du temps. Elle peut donc être très supérieure à la vitesse de la lumière, sans que la causalité relativiste ne soit violée puisque la célérité n'est pas une vitesse à proprement parler.

Nous conclurons ce paragraphe par une remarque importante pour la dynamique. Si une horloge est immobile dans $\mathcal{R}$, le temps propre mesuré entre deux événements coïncide avec celui du référentiel. En revanche, si elle est en mouvement, le temps propre entre les deux mêmes événements est toujours inférieur au temps du référentiel (et ce quelle que soit la forme ou la loi horaire de la trajectoire). On en déduit donc que l'intégrale du temps propre entre deux événements est maximale pour une horloge immobile, une propriété qui nous sera fort utile pour formuler un principe variationnel pour la dynamique relativiste.

[7] Il faut aussi corriger un effet de "red shift" gravitationnel qui n'est descriptible que dans le cadre de la relativité générale. La fréquence de l'horloge est affectée en effet par le champ de pesanteur terrestre, légèrement diminué en vol.

1.4 Transformation de Lorentz

Après cette approche très qualitative, qui nous a permis de comprendre certaines caractéristiques essentielles de la nouvelle cinématique; il nous reste à donner la forme explicite de la transformation de Lorentz, décrivant un changement de référentiel. Nous allons, en fait, oublier pour un temps tout ce que nous avons appris dans les paragraphes précédents et essayer de construire toutes les transformations obéissant à un certain nombre de symétries fondamentales, telles que l'isotropie de l'espace ou l'invariance par translation dans le temps. Nous verrons qu'il n'y a en fait que quatre formes possibles pour une telle transformation. Deux d'entre elles sont inacceptables parce qu'elles conduiraient à abandonner le principe de causalité. Les deux dernières sont la transformation de Galilée, que nous rejetterons également car elle n'obéit pas au principe de relativité, et enfin la transformation de Lorentz. Au cours de cette recherche, nous verrons apparaître certaines propriétés essentielles de la transformation de Lorentz que nous discuterons dans le dernier paragraphe de cette section.

1.4.1 Forme de la transformation de Lorentz

Le choix d'axes pour les deux repères est, encore une fois, celui illustré par la figure 1.1. Nous cherchons donc une transformation $\mathcal{L}(u)$ permettant d'exprimer les coordonnées (ct', x', y', z') d'un événement dans $\mathcal{R}'$ en fonction de celles dans $\mathcal{R}$, (ct, x, y, z). Rappelons que u est la projection algébrique de la vitesse de $\mathcal{R}'$ par rapport à $\mathcal{R}$ sur l'axe du mouvement. Notons tout de suite qu'avec nos conventions l'événement $(0, 0, 0, 0)$ dans $\mathcal{R}$ se transforme en l'événement origine $(0, 0, 0, 0)$ dans $\mathcal{R}'$. Nous obtiendrons avec ce choix d'axes la transformation de Lorentz spéciale. Une simple combinaison avec les rotations et symétries nous permettra ensuite d'obtenir le groupe de Lorentz complet, dont le groupe spécial est un sous–groupe, décrivant des changements de référentiels tout à fait quelconques.

Nous imposerons d'abord à $\mathcal{L}$ d'être une transformation linéaire, homogène. L'invariance de la physique dans une translation arbitraire de l'espace ou du temps impose cette linéarité.

L'ensemble des transformations de Lorentz, paramétrées par la vitesse relative u, doit former un groupe. Considérons en effet trois référentiels: $\mathcal{R}$, $\mathcal{R}'$, en mouvement à la vitesse u par rapport à $\mathcal{R}$, et $\mathcal{R}''$, en mouvement à la vitesse v par rapport à $\mathcal{R}'$ et w par rapport à $\mathcal{R}$ (comme nous avons abandonné le cadre de la relativité galiléenne, w n'est pas égal à $u + v$). La transformation de $\mathcal{R}$ vers $\mathcal{R}'$ peut s'écrire $\mathcal{L}(w)$ ou $\mathcal{L}(v)\mathcal{L}(u)$ (ce produit étant à comprendre comme la composition de deux applications linéaires et donc étant lu de droite à gauche). Le produit de deux transformations de Lorentz définit donc une application de composition interne qui possède évidemment toutes les propriétés d'une loi de groupe. Il existe un élément neutre, l'identité, correspondant au passage d'un référentiel à lui-même et donc à la vitesse nulle. Chaque élément possède un inverse. Il doit en être ainsi, pour qu'à tout événement dans $\mathcal{R}$ corresponde un seul jeu de coordonnées dans $\mathcal{R}'$. La transformation inverse est celle qui donne les coordonnées dans $\mathcal{R}$ en fonction de celles dans $\mathcal{R}'$. La vitesse de l'origine O dans $\mathcal{R}'$ doit bien sûr être $-u$. Si la vitesse de O par rapport à O' n'était pas opposée à la vitesse de O' par rapport à O, nous aurions certainement brisé le principe de relativité. La transformation inverse de $\mathcal{L}(u)$ doit donc être la transformation de Lorentz correspondant à la vitesse $-u$, qui est celle de $\mathcal{R}$ mesurée dans $\mathcal{R}'$: on doit avoir $\mathcal{L}(u)^{-1} = \mathcal{L}(-u)$.

Il n'est pas évident a priori que ce groupe, que nous appellerons "groupe de Lorentz", soit commutatif. En fait, il est possible de montrer que tout groupe paramétré par un paramètre unique, à condition que ce paramétrage soit "suffisamment" continu et dérivable, est isomorphe au groupe additif des réels[8]. Il en résulte immédiatement que tous ces groupes sont abéliens (ou commutatifs). Notons que cet isomorphisme indique qu'on peut, par un changement de variable adéquat, paramétrer

[8]Nous ne démontrerons pas ici cette propriété. On en trouvera une démonstration très élémentaire dans J.M. Lévy–Leblond et al., Am. Journal of Physics, **47**, 1045 (1980). On peut donner des exemples simples de cette propriété. Le groupe multiplicatif des réels (paramétré par la valeur de l'élément) admet une représentation additive évidente qui n'est autre que le logarithme Népérien. Le groupe des rotations autour d'un point, paramétré par l'angle de rotation, est directement paramétré sous forme additive.

le groupe par un paramètre additif, déterminé de façon univoque à un facteur près (à un choix d'unités près). On pourrait ainsi trouver un paramètre $\phi(u)$ tel que la composition de deux transformations de Lorentz s'écrive $\mathcal{L}(\phi(u)+\phi(v)) = \mathcal{L}(\phi(v))\mathcal{L}(\phi(u))$. Nous verrons par la suite que ce paramétrage additif a une signification physique très claire.

Penchons nous d'abord sur les lois de transformation des coordonnées y et z. Comme les axes Ox et $O'x'$ coïncident à tout instant, le fait que $y = z = 0$ implique que $y' = z' = 0$ pour toutes les valeurs de x et t. Ces deux dernières coordonnées ne peuvent donc intervenir dans les lois de transformation de y et z, qui se résument donc à :

$$y' = ay + bz \tag{1.16}$$
$$z' = b'y + a'z \ . \tag{1.17}$$

Les axes Oy et $O'y'$ doivent coïncider à $t = 0$. Si b et b' n'étaient pas nuls, l'axe $O'y'$ correspondrait à des valeurs simultanément non nulles de y et z et ne pourrait donc coïncider avec Oy. La transformation se réduit donc à un simple facteur d'échelle sur y et z. L'isotropie de l'espace impose de plus que les facteurs affectant y et z soient identiques. On a donc simplement:

$$y' = ay \qquad z' = az \ . \tag{1.18}$$

Montrons maintenant que ce facteur a vaut nécessairement 1. Nous avons montré effectivement que $\mathcal{L}(u)^{-1} = \mathcal{L}(-u)$. La transformation inverse est donc décrite par le facteur $a(-u)$, mais aussi par le facteur $1/a(u)$. L'isotropie de l'espace impose de plus que le facteur a ne dépende pas de l'orientation de la vitesse par rapport à l'axe Oy. On a donc $a = 1/a$ et $a = \pm 1$. Si le choix des orientations des axes dans les deux référentiels est cohérent, on a donc finalement $a = 1$. Nous avons montré que la transformation de Lorentz laisse invariantes les coordonnées perpendiculaires à la vitesse relative.

Intéressons nous maintenant à la transformation de x et ct. La transformation la plus générale ferait intervenir les coordonnées y et z. L'invariance par translation perpendiculaire à l'axe des x impose évidemment que y et z n'interviennent pas dans la loi de transformation de x. De même, à x et ct donnés, le temps ct' ne doit pas dépendre de y ou z. Finalement, on peut exprimer la loi de transformation la plus générale par une relation matricielle 2×2:

$$\begin{pmatrix} ct' \\ x' \end{pmatrix} = \begin{pmatrix} a(u) & b(u) \\ e(u) & f(u) \end{pmatrix} \begin{pmatrix} ct \\ x \end{pmatrix} \ . \tag{1.19}$$

On peut préciser considérablement la forme de cette transformation par un simple argument de symétrie. Considérons en effet dans le référentiel $\mathcal{R}'$ un axe $O'X'$ confondu avec, mais d'orientation contraire à, $O'x'$. En un mot, $X' = -x'$. Considérons de même l'axe OX opposé avec l'axe Ox, avec $X = -x$. On peut considérer la transformation donnant X et ct en fonction de X' et ct'. Elle correspond au même changement de référentiel que celui que nous étudions. En effet, dans les deux cas, la vitesse du nouveau référentiel selon l'axe des x ou X est u. La vitesse de O est en effet $-u$ sur $O'x'$ et donc u sur $O'X'$. On en déduit que:

$$\begin{pmatrix} ct \\ X \end{pmatrix} = \begin{pmatrix} a & b \\ e & f \end{pmatrix} \begin{pmatrix} ct' \\ X' \end{pmatrix} \ , \tag{1.20}$$

qu'on peut mettre sous la forme:

$$\begin{pmatrix} ct \\ x \end{pmatrix} = \begin{pmatrix} a & -b \\ -e & f \end{pmatrix} \begin{pmatrix} ct' \\ x' \end{pmatrix} \ . \tag{1.21}$$

Mais cette dernière relation est aussi la transformée inverse de la transformation cherchée:

$$\begin{pmatrix} ct \\ x \end{pmatrix} = \frac{1}{af - be} \begin{pmatrix} f & -b \\ -e & a \end{pmatrix} \begin{pmatrix} ct' \\ x' \end{pmatrix} \ . \tag{1.22}$$

De la comparaison de ces deux formules, on tire immédiatement que le déterminant de la transformation doit être égal à un:

$$af - be = 1 \tag{1.23}$$

et que

$$a = f = \gamma \tag{1.24}$$

(nous changeons un peu les notations pour évoluer vers la forme standard de la transformation de Lorentz).

Nous pouvons préciser encore la forme de la transformation en utilisant le caractère abélien du groupe de Lorentz spécial. En écrivant simplement que $\mathcal{L}(w) = \mathcal{L}(u)\mathcal{L}(v) = \mathcal{L}(v)\mathcal{L}(u)$, on trouve que:

$$\begin{aligned} \gamma(w) &= \gamma(u)\gamma(v) + e(u)b(v) \\ &= \gamma(v)\gamma(u) + e(v)b(u) \ , \end{aligned} \tag{1.25}$$

ce qui ne peut être vérifié pour deux vitesses arbitraires que si $e(u)/b(u)$ est une constante, à moins qu'une de ces fonctions ne s'annule identiquement. Si ce rapport est une constante, un choix convenable d'unités d'espace et de temps permet d'amener sa valeur à ± 1. Nous aurons donc à distinguer 4 cas:

- $e(u) = -b(u)$. La matrice est donc antisymétrique et ses deux coefficients vérifient $\gamma^2 + b^2 = 1$. On peut donc poser $\gamma = \cos\theta$ et $b = \sin\theta$. La matrice de transformation s'écrit donc:

$$\begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix} \ . \tag{1.26}$$

  C'est une simple rotation autour de l'origine dans l'espace–temps. La représentation additive de ce groupe est l'angle de rotation θ.

- $b(u) = 0$. La valeur du déterminant impose alors $\gamma = 1$ et la matrice de la transformation s'écrit:

$$\begin{pmatrix} 1 & 0 \\ e(u) & 1 \end{pmatrix} \ . \tag{1.27}$$

  Ecrire ensuite que le mouvement de O' s'effectue à la vitesse u, c'est à dire que $x = 0$ implique $x' = -ut'$, fixe $e(u) = -\beta = -u/c$. Ce groupe est simplement celui de Galilée dont le paramètre additif est la vitesse u ou la vitesse réduite $\beta = u/c$.

- $e(u) = 0$. Là encore, on doit avoir $\gamma = 1$ et la matrice s'écrit:

$$\begin{pmatrix} 1 & b(u) \\ 0 & 1 \end{pmatrix} \ . \tag{1.28}$$

  Ce groupe, qui diffère du groupe de Galilée en ce qu'il transforme le temps et non l'espace, est le groupe de Caroll. Il est tout naturel de l'éliminer a priori.

- $e(u) = b(u)$. On a alors $\gamma^2 - b^2 = 1$ et on peut poser: $\gamma = \cosh\phi(u)$ et $b(u) = -\sinh\phi(u)$ (la raison de ce choix de signe apparaîtra clairement plus tard). La matrice de transformation:

$$\begin{pmatrix} \cosh\phi & -\sinh\phi \\ -\sinh\phi & \cosh\phi \end{pmatrix} \tag{1.29}$$

  est alors simplement celle d'une rotation hyperbolique (rotation autour de l'origine d'un angle imaginaire pur). La représentation additive de ce groupe (que nous appellerons "groupe de Lorentz", en faisant fi du suspense) est simplement l'angle de rotation ϕ, dont nous donnerons dans un moment l'interprétation physique.

Nous n'avons utilisé jusque là que des arguments très généraux de symétrie et de réciprocité. Il est déjà très remarquable que nous n'ayons plus le choix qu'entre quatre groupes, chacun paramétré par un seul paramètre additif. Pour choisir la forme correcte, nous pouvons employer deux arguments supplémentaires[9].

Si nous voulons que la causalité soit une notion indépendante du référentiel, il faut au moins exiger de la transformation qu'elle préserve l'ordre temporel des événements (au moins pour certains couples d'événements, ceux qui sont dans le cône de lumière l'un de l'autre). Il doit donc exister des classes d'événements pour lesquelles le signe de t ne doit changer dans aucune transformation. Il est facile de voir que le groupe des rotations spatio–temporelles et le groupe de Caroll contiennent des transformations changeant le signe de t pour tout événement. Ils sont donc exclus par de simples exigences de causalité.

Il ne nous reste que le groupe de Galilée et le groupe de Lorentz. Le premier, qui conduit à la composition des vitesses au sens ordinaire, n'est pas acceptable. La transformation cherchée doit appartenir donc au groupe de Lorentz (à celui des rotations hyperboliques). Il ne nous reste plus qu'à préciser la valeur du paramètre additif $\phi(u)$, que nous appellerons "rapidité". Il suffit pour cela d'écrire que O est animé, dans $\mathcal{R}'$, d'un mouvement uniforme à la vitesse $-u$ ($x' = -ut'$). On déduit immédiatement de la transformation que

$$x' = -\sinh\phi(u)ct, \quad ct' = \cosh\phi(u)ct \tag{1.30}$$

et

$$\tanh\phi(u) = \frac{u}{c} = \beta \ . \tag{1.31}$$

soit encore

$$\sinh\phi(u) = \gamma\beta \ , \qquad \gamma = \cosh\phi(u) = \frac{1}{\sqrt{1-\beta^2}} \ . \tag{1.32}$$

Notons que le facteur γ peut prendre des valeurs arbitrairement grandes. La rapidité peut donc aussi être arbitrairement grande. Si la vitesse u est limitée par la vitesse de la lumière, il n'en est pas de même pour le paramètre "naturel" du groupe de Lorentz.

Avec ces valeurs, nous avons complètement déterminé la transformation de Lorentz, qui peut s'écrire:

$$\begin{pmatrix} ct' \\ x' \end{pmatrix} = \begin{pmatrix} \gamma & -\gamma\beta \\ -\gamma\beta & \gamma \end{pmatrix} \begin{pmatrix} ct \\ x \end{pmatrix} \ , \tag{1.33}$$

la transformation inverse étant évidemment donnée par

$$\begin{pmatrix} ct \\ x \end{pmatrix} = \begin{pmatrix} \gamma & \gamma\beta \\ \gamma\beta & \gamma \end{pmatrix} \begin{pmatrix} ct' \\ x' \end{pmatrix} \ , \tag{1.34}$$

(il suffit de changer le signe de la vitesse relative). En termes de coordonnées et de temps, on peut aussi écrire la transformation directe sous la forme:

$$x' = \gamma(x - ut) \tag{1.35}$$

$$t' = \gamma\left(t - \frac{ux}{c^2}\right) \tag{1.36}$$

et la transformation inverse sous la forme:

$$x = \gamma(x' + ut') \tag{1.37}$$

$$t = \gamma\left(t' + \frac{ux'}{c^2}\right) \tag{1.38}$$

[9] A ce point, imposer la constance de la vitesse de la lumière ou l'invariance de l'intervalle suffirait à choisir le dernier groupe. Nous allons suivre une démarche un plus détaillée pour montrer que la sélection peut aller encore plus loin sans le postulat de relativité.

A toutes les relations précédentes, il convient bien sûr d'ajouter l'invariance des coordonnées transverses y et z. Notons immédiatement que, si on ne retient dans la transformation de Lorentz que les termes du premier ordre en u/c, on retrouve la transformation de Galilée. La cinématique classique apparaît donc bien comme une limite de la cinématique relativiste pour des vitesses d'entraînement faibles devant celle de la lumière.

Notons que nous avons fait, dans tout ce paragraphe, un choix d'axes bien particulier. Comme il nous a conduit à une forme univoque de la transformation de Lorentz, nous n'avons pas restreint la généralité. On peut avoir cependant à composer des transformations de Lorentz correspondant à des directions de vitesses différentes. La transformation de Lorentz s'écrira alors $\mathcal{L}(\mathbf{u})$, où $\mathbf{u}$ est le vecteur vitesse de $\mathcal{R}'$ par rapport à $\mathcal{R}$, de direction arbitraire. Pour écrire ce genre de changement de référentiel, il convient de composer la transformation que nous venons d'écrire avec les rotations arbitraires d'espace, avec les réflexions d'espace et même, éventuellement, avec les réflexions du temps. On obtient ainsi le "groupe de Lorentz complet", qui décrit tous les changements de référentiels. On distingue parfois, à l'intérieur du groupe complet , plusieurs sous–groupes:

- Le groupe de Lorentz "propre" comprenant la transformation de Lorentz combinée avec les rotations spatiales. Sauf cas spécial, il suffit à décrire un changement de référentiel avec une direction de vitesse arbitraire. Le déterminant de la matrice correspondante est 1.
- Le groupe de Lorentz "orthochrome" contient le groupe de Lorentz combiné éventuellement avec des réflexions d'espace (nous avions exclu explicitement ces transformations dans notre discussion en imposant aux directions des axes d'être consistantes). Le déterminant de la matrice peut alors être ± 1.

1.5 Propriétés de la transformation de Lorentz

Avant d'appliquer la transformation de Lorentz à des situations physiques, nous allons nous pencher plus en détails sur certaines de ses propriétés. La première, qui découle directement de l'expression de la transformation comme une rotation hyperbolique, est que la transformation de Lorentz conserve l'intervalle. Nous aurions en fait pu prendre cette hypothèse comme point de départ et construire à partir de là la transformation.

Un point essentiel de ce paragraphe sera de comprendre qu'il y a trois quantités fondamentales différentes décrivant la vitesse d'un référentiel par rapport à un autre, correspondant à trois situations expérimentales différentes pour déterminer cette vitesse.

1.5.1 Composition des transformations

Revenons à la situation déjà évoquée de trois référentiels en mouvement relatif. La loi de composition des rapidités nous permet d'écrire de manière évidente:

$$\phi(w) = \phi(u) + \phi(v) \ . \tag{1.39}$$

Nous pouvons en tirer facilement la loi de composition des vitesses relativistes, sous une forme simplifiée correspondant à des vitesses qui sont toutes colinéaires (nous généraliserons au paragraphe suivant), en écrivant w en fonction de u et v. Il suffit pour cela de prendre le cosh et le sinh de l'équation précédente. On obtient alors:

$$\sinh\phi(w) = \beta(w)\gamma(w) = \gamma(u)\gamma(v)(\beta(u)+\beta(v)) \tag{1.40}$$

$$\cosh\phi(w) = \gamma(w) = \gamma(u)\gamma(v)(1+\beta(u)\beta(v)) \ , \tag{1.41}$$

d'où on tire immédiatement:

$$\beta(w) = \frac{\beta(u)+\beta(v)}{1+\beta(u)\beta(v)} \ , \tag{1.42}$$

ou encore

$$w = \frac{u+v}{1+uv/c^2} \ , \tag{1.43}$$

Cette loi remplace la simple addition des vitesses de la relativité galiléenne. Notons là encore qu'on retrouve la loi galiléenne d'addition des vitesse pour la composition de vitesses toutes deux petites devant la vitesse de la lumière. Remarquons également que cette loi prédit correctement l'invariance de la vitesse de la lumière: si $\beta(u) = 1$ ou si $\beta(v) = 1$, on trouve immédiatement $\beta(w) = 1$. Cela montre aussi qu'on ne peut dépasser la vitesse de la lumière en courant dans le couloir d'un train se déplaçant à une vitesse proche de c. Encore une fois, si l'accumulation de vitesses ne peut conduire à une vitesse supérieure à celle de la lumière, les rapidités s'accumulent sans limite. Il est d'ailleurs fructueux à ce point de comparer ces notions de vitesses et de rapidité.

1.5.2 Vitesse, célérité et rapidité

Nous sommes maintenant en possession de trois quantités différentes décrivant le mouvement d'un référentiel par rapport à un autre. Il n'est que temps d'examiner les différences entre ces quantités et de préciser leur sens physique.

La première définition de la vitesse relative de $\mathcal{R}'$ (le référentiel du contrôleur, pour reprendre le vocabulaire ferroviaire) par rapport à $\mathcal{R}$ (le chef de gare) est la vitesse u, vitesse de l'origine O' dans $\mathcal{R}$. Imaginons, pour bien insister sur le sens physique de ces quantités, que le contrôleur désire déterminer sa vitesse. La première méthode est de repérer, sur les horloges de deux gares successives, ses temps de passage. La liste des tarifs, imprimée dans $\mathcal{R}$, lui donnant la distance (mesurée dans $\mathcal{R}$) entre ces gares, il en déduira sa vitesse. Cette vitesse, mesurée dans l'espace de $\mathcal{R}$ avec le temps de $\mathcal{R}$, ou dans l'espace de $\mathcal{R}'$ avec le temps de $\mathcal{R}'$, est bien entendu la vitesse u.

L'autre définition que nous avons déjà rencontrée est celle de la célérité. Rappelons que c'est la définition qui intéresse le voyageur interstellaire, puisqu'elle mesure la distance parcourue dans le référentiel fixe par unité de temps du mobile. Nous avons vu que la célérité s'exprimait par γu ou $\gamma\beta = \sinh\phi$ en unités réduites. Cette vitesse est celle que déterminerait le contrôleur en mesurant la durée qui s'écoule entre les deux gares avec sa propre montre. Notons enfin que célérité et vitesse sont identiques en relativité galiléenne, en raison de l'universalité du temps.

Nous sommes maintenant en possession d'une troisième "définition" de la vitesse en termes de la rapidité. Si nous savons déjà que son intérêt réside dans son caractère additif, à quelle expérience correspondrait-elle?

Le contrôleur dispose d'une troisième méthode pour déterminer sa vitesse, même si les vitres sont occultées. Supposons qu'il soit initialement immobile dans la première gare mais qu'il dispose d'un accéléromètre (un simple fil à plomb lui suffirait). Cet accéléromètre mesure la variation de la vitesse du train par unité de temps du train (ce temps est un temps propre). En intégrant les indications de l'accéléromètre sur toute la phase d'accélération, le contrôleur pourra déterminer sa vitesse finale.

On peut traiter ce cas très simple de cinématique d'un mouvement accéléré sans recourir à la relativité générale. Le référentiel du train n'est plus galiléen. Nous considérerons en revanche comme référentiel $\mathcal{R}'$ le référentiel tangent au mouvement pour une valeur du temps propre τ (obtenu, rappelons–le, par intégration des temps propres dans les référentiels tangents successifs). A un instant donné, la vitesse de ce référentiel par rapport à $\mathcal{R}$ est v, qui passe de 0 à u pendant la phase d'accélération. Pendant un intervalle de temps infinitésimal $d\tau$, la vitesse du train dans le référentiel tangent passe de 0 à dv'. L'accélération mesurée dans le train (ou plutôt dans le référentiel tangent) vaut donc $a = dv'/d\tau$. Pendant l'intervalle de temps correspondant, la vitesse dans $\mathcal{R}$ passe de v à $v + dv$. En utilisant la loi de composition des vitesses du paragraphe précédent, nous pouvons écrire l'accroissement de vitesse dv' dans $\mathcal{R}'$ en fonction de la nouvelle vitesse dans $\mathcal{R}$ $(v+dv)$ et de la vitesse v de $\mathcal{R}'$ par rapport à $\mathcal{R}$:

$$dv' = \frac{v+dv-v}{1-v(v+dv)/c^2} \ , \tag{1.44}$$

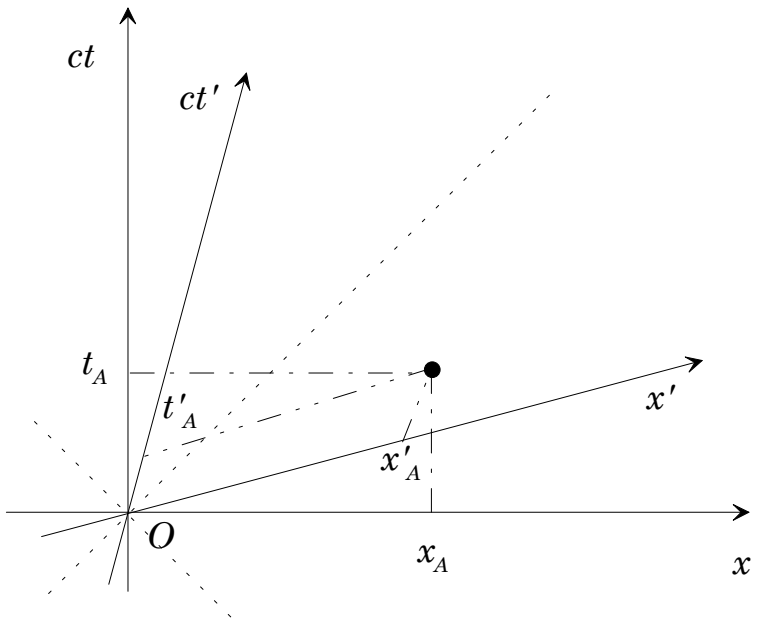

Figure 1.5: La transformation de Lorentz interprétée géométriquement comme une rotation hyperbolique dans l'espace temps. Ses directions propres coïncident avec le cône de lumière de l'origine.

soit encore

$$dv' = \frac{dv}{1 - v^2/c^2} \ . \tag{1.45}$$

La vitesse finale mesurée par cette méthode sera donc:

$$\int a d\tau = \int \frac{dv}{1 - v^2/c^2} = c \ \mathrm{arctanh} \frac{u}{c} \ . \tag{1.46}$$

Elle coïncide donc avec la définition de la rapidité, qui prend ainsi un sens physique très fort. Cette définition en termes d'accélération accumulée nous fait comprendre pourquoi la rapidité n'est pas bornée. On peut en effet avoir un mouvement indéfiniment accéléré en relativité restreinte (nous écrirons ce mouvement pour décrire celui de particules chargées soumises à une force constante dans un champ électrique uniforme). La vitesse tend asymptotiquement vers c, l'accélération habituelle (mesurée dans le référentiel fixe) tend vers zéro mais l'accélération mesurée comme ci-dessus demeure constante et la rapidité s'accumule. Notons pour finir qu'en relativité galiléenne la rapidité et la vitesse coïncident aussi.

1.5.3 Géométrie de la transformation de Lorentz

Penchons nous maintenant sur l'interprétation géométrique de la transformation de Lorentz dans l'espace temps. Si une rotation ordinaire autour de l'origine est bien connue, il n'en est pas nécessairement de même pour une rotation hyperbolique. La figure 1.5 illustre la géométrie de cette transformation. Au contraire d'une rotation, elle ne conserve pas l'angle entre les axes. Pour $u > 0$, les deux nouveaux axes sont à l'intérieur du premier quadrant. On a illustré sur la figure 1.5 un événement (par un point), ainsi que ses nouvelles et ses anciennes coordonnées, obtenues par projection sur les axes correspondants.

On peut préciser encore notre interprétation de cette transformation, en considérant ses valeurs propres et vecteurs propres. L'équation caractéristique s'écrit $\lambda^2 - 2\gamma\lambda + 1 = 0$. Elle admet donc deux

valeurs propres réelles inverses l'une de l'autre:

$$\frac{1 \pm \beta}{\sqrt{1-\beta^2}} = \frac{1}{\sqrt{1 \mp \beta}} = \exp \pm \phi \ . \tag{1.47}$$

Les vecteurs propres s'obtiennent sans difficulté. Ils correspondent à $x = ct$ (pour la valeur propre supérieure à un) et $x = -ct$. Les directions propres de la transformation de Lorentz ne sont autres que celles du cône de lumière du point O (directions que nous avons également représentées sur la figure 1.5).

1.6 Conséquences de la transformation de Lorentz

1.6.1 Retour sur nos expériences de pensée

Nous reviendrons brièvement, dans ce paragraphe, sur les deux expériences de pensée que nous avons discutées au début de ce chapitre. Nous allons les décrire en termes d'événements et utiliser la transformation de Lorentz complète pour exprimer les changements de référentiels. Nous pourrons en particulier préciser la valeur de la non–simultanéité dans la seconde expérience, ce que nous ne pouvions faire sans la transformation.

Pour la première expérience, il nous faut distinguer trois événements: le départ de l'impulsion de O' ($x'_1 = y'_1 = t'_1 = 0$) (nous n'écrirons jamais la coordonnée z qui ne joue aucun rôle); la réflexion sur le miroir de coordonnées $x'_2 = 0, y'_2 = L, t'_2 = L/c$, et le retour de l'impulsion en O' de coordonnées $x'_3 = y'_3 = 0, t'_3 = 2L/c$. La transformation de Lorentz donne les positions de ces trois événements dans $\mathcal{R}$:

$$x_1 = y_1 = t_1 = 0 \tag{1.48}$$

$$x_2 = \gamma u L/c, \ y_2 = L, \ t_2 = \gamma L/c \tag{1.49}$$

$$x_3 = 2\gamma u L/c, \ y_3 = 0, \ t_3 = 2\gamma L/c \ , \tag{1.50}$$

qui nous redonne immédiatement la dilatation des temps.

Pour la seconde expérience, nous n'avons à considérer que l'allumage des feux. La réception des signaux coïncide en effet avec l'événement origine. On a , dans $\mathcal{R}$:

$$x_A = -L \qquad t_A = -L/c \tag{1.51}$$

$$x_B = L \qquad t_B = -L/c \ . \tag{1.52}$$

On en déduit immédiatement, dans $\mathcal{R}'$,

$$x'_A = -\gamma L(1-\beta) \qquad t'_A = -\gamma(1-\beta)L/c \tag{1.53}$$

$$x'_B = \gamma L(1+\beta) \qquad t'_B = -\gamma(1+\beta)L/c \ , \tag{1.54}$$

ce qui montre clairement que les instants d'allumage dans $\mathcal{R}'$ différent de $2\gamma\beta L/c$.

1.6.2 Intervalles et simultanéité

Nous allons établir deux propriétés qui nous permettront de revenir sur la notion de causalité relativiste.

Si deux événements sont séparés par un intervalle de genre temps, il existe un référentiel dans lequel ils se produisent au même endroit.

Considérons donc deux événements dans un référentiel quelconque $\mathcal{R}$. On peut toujours choisir l'un comme origine et faire en sorte, par un choix d'axes, que l'autre se produise sur l'axe Ox. Leurs

coordonnées sont alors $(0,0)$ et (ct,x) . On a $|ct| > |x|$ puisque leur intervalle est du genre temps. Prenons un nouveau référentiel $\mathcal{R}'$. Dans ce référentiel, les deux événements se produisent au même endroit si $x' = \gamma(x - ut) = 0$. Il suffit pour cela que $u = x/t$, qui est bien inférieur à c. Notons que le carré de l'intervalle est alors simplement le carré du temps propre, du temps qui s'écoule entre les deux événements dans le référentiel où ils se produisent au même point.

Si les deux événements se produisent au même endroit dans ce référentiel, il peuvent être liés par un lien de causalité. Une autre manière de voir cette propriété est d'imaginer un signal se propageant d'un événement à l'autre. Comme l'intervalle est du genre temps, ce signal se propage moins vite que la lumière. On peut donc lui associer un référentiel qui n'est, d'ailleurs, autre que $\mathcal{R}'$. On peut noter enfin que l'instant t' auquel se produit le second événement dans $\mathcal{R}'$ est positif si et seulement si t est positif. Si deux événements se produisent au même point dans un référentiel, l'ordre temporel des événements n'est modifié par aucune transformation de Lorentz. En un mot, la notion de causalité est complètement préservée par les changements de référentiels.

Si deux événements sont séparés par un intervalle du genre espace, il existe un référentiel où ils se produisent simultanément.

Nous emploierons les mêmes notations. Cette fois, $|ct| < |x|$. Dans $\mathcal{R}'$ les deux événements se produisent au même instant si $t' = \gamma(t - ux/c^2) = 0$, c'est à dire si $u = c^2t/x$, qui est, là encore, plus petit que c. Le fait que les événements se produisent au même instant à des endroits différents prouve qu'ils ne peuvent être liés par un lien de causalité, la relativité n'admettant pas d'action instantanée à distance. On peut voir rapidement que, dans ce cas, le signe de t dépend de la transformation. Comme il n'y a pas de relation de causalité relativiste entre ces événements, leur ordre temporel peut dépendre du référentiel. Notons enfin que le carré de l'intervalle est dans ce cas l'opposé de la distance entre les deux événements dans le référentiel où ils se produisent au même instant. Il s'agit d'une longueur propre, telle que nous la définirons plus précisément dans un instant.

1.6.3 Loi de composition des vitesses

Nous généraliserons dans ce paragraphe la loi de composition des transformations de Lorentz qui nous a déjà permis, au paragraphe précédent de traiter le cas de deux vitesses colinéaires. Nous considérerons ici une situation physique légèrement différente. Un mobile est en mouvement avec une vitesse $\mathbf{v}$ dans le référentiel $\mathcal{R}'$, entraîné à une vitesse u (selon Ox) par rapport au référentiel $\mathcal{R}$. En écrivant les accroissements infinitésimaux dx', dy', dz' et dt' dans $\mathcal{R}'$ et en utilisant la transformation de Lorentz, on en déduit les accroissements correspondants dans $\mathcal{R}$:

$$dx = \gamma(dx' + udt') \tag{1.55}$$
$$dt = \gamma(dt' + udx'/c^2) \tag{1.56}$$
$$dy = dy' \tag{1.57}$$
$$dz = dz' \ . \tag{1.58}$$

On peut alors calculer sans difficultés les vitesses dans $\mathcal{R}$. On obtient:

$$v_x = \frac{v'_x + u}{1 + uv'_x/c^2} \tag{1.59}$$
$$v_y = \frac{v'_y}{\gamma(1 + uv'_x/c^2)} \ , \tag{1.60}$$

la transformation inverse s'obtenant trivialement en changeant le signe de u dans les expressions précédentes. On vérifiera, à titre d'exercice, que la vitesse de la lumière est bien invariante dans cette transformation.

Aberration des étoiles

Nous étudierons ici une application immédiate de la loi de composition des vitesses. La lumière provenant d'une étoile apparaît déviée par la composition de sa vitesse avec celle du référentiel terrestre par rapport au système solaire. La position apparente d'une étoile dans le ciel dépend donc dans une petite mesure de la position de la terre sur son orbite (chaque étoile semble décrire une petite ellipse annuelle autour de sa position moyenne). Pour simplifier la géométrie, nous considérerons le référentiel $\mathcal{R}$ comme celui lié au système solaire. La lumière de l'étoile arrive parallèlement à l'axe Oy. La vitesse de cette lumière est donc $v_y = -c$.

Le référentiel $\mathcal{R}'$ est celui de la terre, entraîné à la vitesse u selon Ox. De la loi de composition des vitesses, on déduit les composantes de la vitesse de la lumière de cette étoile dans le référentiel $\mathcal{R}'$:

$$v'_x = -u \tag{1.61}$$

$$v'_y = -\frac{c}{\gamma} \tag{1.62}$$

On "voit" donc dans $\mathcal{R}'$ la lumière de l'étoile provenir d'une direction inclinée par rapport à l'axe $O'y'$, d'un angle $\theta = \arctan \beta\gamma$. L'application de la cinématique galiléenne prévoit aussi une déviation apparente mais elle n'est, comme on le vérifiera aisément, que $\theta = \arctan \beta$. La différence entre ces deux quantités est mesurable et son observation fut une belle confirmation de la relativité restreinte.

Dans le même genre de situation, nous aurions pu nous intéresser aussi à la fréquence de la lumière reçue. Nous l'aurions trouvée modifiée, d'une quantité différente de l'effet Doppler classique (il existe en particulier un effet Doppler du second ordre pour une vitesse d'entraînement perpendiculaire à la propagation). Nous pourrons traiter ce problème plus efficacement quand nous aurons introduit, au chapitre suivant, les notations tensorielles.

1.6.4 Contraction des longueurs

Dans cette nouvelle conséquence immédiate de la transformation de Lorentz, nous allons retrouver l'hypothèse heuristique que Lorentz avait introduite pour expliquer le résultat négatif de l'expérience de Michelson.

Le problème que nous abordons ici est celui de la définition de la longueur d'un objet en mouvement. Pour fixer les idées, nous considérons une règle rigide, de longueur L', immobile dans le référentiel mobile $\mathcal{R}'$, confondue avec l'axe $O'x'$ (et ayant une extrémité en O'). Comment des observateurs de $\mathcal{R}$ peuvent-ils déterminer la longueur de cette règle? Il y a deux stratégies.

L'observateur O peut d'abord déterminer la vitesse u de la règle, par exemple par vélocimétrie Doppler. Il peut alors mesurer la durée Δt pendant laquelle la règle défile devant lui. Il en déduira alors sa longueur $L = u\Delta t$. On peut aussi prendre une photographie instantanée de la règle. En fait, on peut repérer, à un instant t donné, les observateurs de $\mathcal{R}$ qui sont en face des extrémités de la règle et mesurer leur distance. A titre d'exercice, nous allons examiner ces deux procédures et montrer qu'elles fournissent la même longueur.

Dans la première méthode, les deux événements A et B à considérer sont le passage en O des deux extrémités de la règle. Les coordonnées de ces deux événements dans $\mathcal{R}'$ sont

$$x'_A = t'_A = 0 \tag{1.63}$$

$$x'_B = -L' \qquad t'_B = L'/u \tag{1.64}$$

(par convention, l'extrémité O' de la règle passe d'abord devant O). Leurs coordonnées dans $\mathcal{R}$ sont alors:

$$x_A = t_A = 0 \tag{1.65}$$

$$x_B = 0 \qquad t_B = \gamma\frac{L'}{u}\left(1 - \frac{u^2}{c^2}\right) = \frac{L'}{\gamma u}\ . \tag{1.66}$$

On retrouve évidemment que ces deux événements se produisent en O. La durée de passage Δt est égale à t_B, et la longueur L de la règle ainsi mesurée est:

$$L = \frac{L'}{\gamma} \ . \tag{1.67}$$

Pour la deuxième méthode, on considère les deux événements A et B représentant les extrémités de la règle à un instant donné dans $\mathcal{R}$. Pour simplifier, nous prendrons l'instant origine. Les coordonnées dans $\mathcal{R}$ de ces événements sont donc:

$$x_A = t_A = 0 \tag{1.68}$$

$$x_B = -L \qquad t_B = 0 \ . \tag{1.69}$$

Leurs coordonnées dans $\mathcal{R}'$ sont donc:

$$x'_A = t'_A = 0 \tag{1.70}$$

$$x'_B = -\gamma L \qquad t'_B = \gamma u L/c^2 \ . \tag{1.71}$$

Comme on doit aussi avoir $x'_B = -L'$, on en déduit encore

$$L = \frac{L'}{\gamma} \ . \tag{1.72}$$

Les deux procédures conduisent donc, heureusement, à la même longueur. Rappelons pour finir que la longueur de la règle n'apparaît pas modifiée quand elle est perpendiculaire à Ox.

Une règle en mouvement dans la direction de sa longueur apparaît donc plus courte que dans un référentiel où elle est au repos. Cette contraction des longueurs est, dans la première méthode, une conséquence directe de la dilatation des temps. Dans la deuxième approche, elle est une conséquence de la non invariance de la simultanéité. Si on regarde les extrémités à un même instant dans $\mathcal{R}$, on les voit à deux instants différents dans $\mathcal{R}'$. Qui dit différence sur les temps, dit légère différence sur les positions.

Une application immédiate de la contraction des longueurs est la définition de la célérité. Reprenons le jumeau voyageur du paradoxe de Langevin. Dans le référentiel fixe, il met 4 ans à atteindre l'étoile la plus proche, à une vitesse proche de celle de la lumière. Dans son référentiel propre, il ne met que $4/\gamma$ ans. Cependant, la vitesse de l'étoile par rapport à lui est égale (en module) à sa vitesse par rapport à la terre et donc proche de c. En revanche, la distance de l'étoile n'est plus que de $4/\gamma$ années lumière. On pourra, à titre d'exercice, écrire complètement le paradoxe des jumeaux en termes d'événements.

Un "Paradoxe" résolu

Pour illustrer encore cette notion de contraction des longueurs, considérons un nouveau problème de trains. Un train de longueur L (référentiel propre $\mathcal{R}'$) entre dans un tunnel (immobile) de longueur L exactement égale. $\mathcal{R}$ est le référentiel du tunnel. Train et tunnel sont alignés avec Ox et $O'x'$. A $t = t' = 0$, l'arrière du train passe juste dans l'entrée du tunnel, située en O (ou O'). Que voient le contrôleur et le chef de gare (en mission d'inspection dans le tunnel)? Pour le chef de gare, la longueur du train en mouvement est inférieure à celle du tunnel et la locomotive sort du tunnel un peu après que le dernier wagon n'y ait pénétré. Pour le contrôleur, en revanche, c'est le tunnel qui est un peu plus court que le train et la locomotive sort du tunnel avant que le dernier wagon n'y entre. La solution de ce paradoxe apparent est bien sûr dans la non universalité de la simultanéité. Avant et après ne sont pas des absolus pour des points situés à des endroits différents (et qui ne sont pas dans le cône de lumière l'un de l'autre).

Pour nous en convaincre, nous écrirons, dans les deux référentiels, les coordonnées des deux événements importants. L'un représente l'entrée du dernier wagon dans le tunnel et coïncide avec l'événement origine dans les deux référentiels. L'autre événement est la sortie de la locomotive, dont les coordonnées dans $\mathcal{R}$ sont:

$$x_2 = L \qquad t_2 = \frac{L}{u}\left(1 - \frac{1}{\gamma}\right) \ . \tag{1.73}$$

On peut obtenir simplement t_2 en disant que la longueur du train dans $\mathcal{R}$ est L/γ. A l'instant 0, la locomotive est donc à une distance $L(1 - 1/\gamma)$ de la sortie et elle parcourt cette distance à la vitesse u. La sortie de la locomotive dans $\mathcal{R}$ se produit donc bien après l'entrée du dernier wagon. On peut vérifier, par un calcul élémentaire, que le carré de l'intervalle entre les événements 1 et 2 est $c^2(1-\gamma)/u^2\gamma^2$, négatif. L'intervalle entre les événements est toujours du genre espace. La notion de passé et de futur pour ces deux événements n'est donc pas nécessairement invariante.

On peut alors écrire les coordonnées correspondantes dans $\mathcal{R}'$. Après un calcul sans difficulté, on trouve:

$$x'_2 = L \ , \tag{1.74}$$

ce qui ne fait guère que vérifier la cohérence du calcul et

$$t'_2 = -\frac{L}{u}\frac{\gamma - 1}{\gamma} \ . \tag{1.75}$$

L'événement 2 (sortie de la locomotive) s'est donc produit, dans $\mathcal{R}'$, avant que l'arrière du train n'entre dans le tunnel, comme nous nous y attendions. Si il y a dans ce problème un paradoxe, il ne concerne pas la validité et la cohérence de la transformation de Lorentz. Ce n'est que la différence entre les prédictions de la relativité sur le temps et notre sens commun qui crée l'apparence paradoxale de ces situations. Ce n'est qu'avec une fréquentation assidue de ce genre de problèmes que peut se développer une intuition relativiste. Le problème est similaire, bien que moins ardu, à celui qu'on rencontre pour se forger une intuition en mécanique quantique.

Chapitre 2

Notations Quadridimensionnelles

Nous allons introduire dans ce chapitre des notions mathématiques permettant de traiter la relativité restreinte d'une façon particulièrement aisée. En nous plaçant dans un espace–temps à quatre dimensions, en le munissant d'une structure d'espace vectoriel, d'un produit scalaire et d'une norme, nous pourrons réinterpréter d'une façon très efficace les notions introduites dans le chapitre précédent. Les objets physiques, indépendants du choix du référentiel, seront alors décrits comme des vecteurs, ou, plus généralement des tenseurs (matrices) dans cet espace. La transformation de Lorentz apparaîtra alors comme un simple changement de base. Les lois de transformation de toutes les quantités physiques se déduiront alors simplement des règles de changement de base. Pour manipuler commodément ces vecteurs et ces tenseurs, nous introduirons également des notations tensorielles très puissantes dues à Einstein. Nous allons commencer par introduire la notion d'espace temps à partir de l'événement.

2.1 4–vecteur position d'un événement.

2.1.1 Coordonnées contravariantes

Un événement est complètement défini dans un référentiel par la donnée des quatre nombres (ct, x, y, z). Nous considérerons donc un espace vectoriel à quatre dimensions (d'où le nom de 4–vecteur) muni d'un base $\mathbf{e}_\mu$ (l'indice μ variant entre 0 et 3) et nous associerons à cet événement le vecteur (ou plutôt le 4–vecteur):

$$\mathbf{R} = \sum_\mu x^\mu \mathbf{e}_\mu \ , \tag{2.1}$$

en posant

$$x^0 = ct, \ x^1 = x, \ x^2 = y, \ x^3 = z \ . \tag{2.2}$$

Nous appellerons les x^μ les composantes contravariantes du 4–vecteur événement. Par convention, dans toute la suite, on utilisera des indices supérieurs pour représenter ces composantes (le risque de confusion avec des exposants est pratiquement nul). Par convention aussi, nous ne préciserons pas le domaine sur lequel s'effectuent les sommations. Un indice représenté par une lettre grecque variera toujours entre 0 et 3 (0 représentant la coordonnée temporelle). Quand nous aurons besoin d'un indice ne parcourant que les valeurs 1 à 3 (et donc associé à de simples composantes spatiales), nous utiliserons un indice romain[1]:

$$x^i = x^{1,2,3} \ . \tag{2.3}$$

L'écriture systématique des signes somme serait extrêmement fastidieuse, surtout quand nous manipulerons, avec les tenseurs, des sommes multiples. Nous adopterons donc la convention d'Einstein

[1] Cette convention est très largement répandue dans les manuels récents. Signalons cependant que le Landau de théorie des champs, dont nous recommandons la lecture, utilise une convention strictement inverse.

de *sommation sur les indices répétés.* Quand, dans une expression, un même indice apparaît en haut et en bas, on doit sommer l'expression sur toutes les valeurs possibles de cet indice (0 à 3 pour un indice grec, 1 à 3 pour un indice romain). On écrira donc simplement:

$$\mathbf{R} = \sum_{\mu} x^{\mu}\mathbf{e}_{\mu} = x^{\mu}\mathbf{e}_{\mu} \ . \tag{2.4}$$

Insistons sur le fait, dont la raison apparaîtra plus clairement plus tard, que l'on ne somme implicitement sur un indice que s'il apparaît une fois en haut et une fois en bas. Le même indice apparaissant deux fois en bas ne doit pas être sommé. Une écriture comme $T_{\mu\mu}$, par exemple, représentera un élément diagonal d'un tenseur (disons à ce stade d'une matrice) et pas une somme. La trace de cette matrice s'écrirait $T_{\mu}{}^{\mu}$, la convention de somme s'appliquant alors à l'indice répété. Nous verrons dans la suite de cet exposé à quel point cette simple convention de sommation implicite améliore les écritures. Nous appellerons indice "libre" un indice sur lequel la règle de sommation ne s'applique pas (qui donc reste en tant qu'indice dans l'expression finale) et "muet" un indice faisant l'objet d'une sommation implicite (et n'apparaissant pas en tant que tel dans l'expression finale[2]). Nous respecterons, pour les indices libres, la règle de "balancement". Dans une équation, les indices libres apparaissant dans les deux membres doivent se correspondre un à un et apparaître en même position (haut ou bas). Nous comprendrons plus tard que, si ce n'était pas le cas, les objets décrits par les deux membres de l'équation ne seraient pas de même nature. Ces règles de sommation et de balancement, qui portent sur la typographie des équations, constituent en fait, en plus d'une simplification notable, un système de garde-fous rendant impossible l'écriture d'expressions absurdes. Elles jouent en relativité le rôle des notations de Dirac de la mécanique quantique qui relient de façon univoque la nature de l'objet (fonction d'onde, objet du dual, produit scalaire, opérateur) à son écriture (ket, bra, bra ket, ket bra).

Nous pouvons munir notre espace–temps d'un produit scalaire. Considérant deux 4–vecteurs associés à deux événements, $\mathbf{R} = x^{\mu}\mathbf{e}_{\mu}$ et $\mathbf{S} = y^{\mu}\mathbf{e}_{\mu}$, nous pouvons écrire le produit scalaire comme

$$\mathbf{R}\cdot\mathbf{S} = x^{\mu}y^{\nu}\mathbf{e}_{\mu}\cdot\mathbf{e}_{\nu} \ . \tag{2.5}$$

Nous souhaiterions bien sûr que notre produit scalaire ait une expression compatible avec celle de l'intervalle. Il faut pour celà que $\mathbf{R}\cdot\mathbf{R} = c^2t^2 - x^2 - y^2 - z^2$. En posant:

$$g_{\mu\nu} = \mathbf{e}_{\mu}\cdot\mathbf{e}_{\nu} \ , \tag{2.6}$$

le produit scalaire s'écrit:

$$\mathbf{R}\cdot\mathbf{S} = g_{\mu\nu}x^{\mu}y^{\nu} \tag{2.7}$$

(où l'on voit bien la simplification apportée par la convention de sommation implicite). Nous aurons un produit scalaire convenable si le tableau de nombres $g_{\mu\nu}$ (que nous ne pouvons encore considérer comme un opérateur ou un tenseur – et que nous nommerons néanmoins le "tenseur métrique") s'écrit:

$$g_{\mu\nu} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} \tag{2.8}$$

Dans ce tableau, μ est l'indice ligne et ν l'indice colonne. Cette convention sur le tenseur métrique peut aussi être vue comme une condition d'orthonormalité pour la base $\mathbf{e}_{\mu}$. Nous ne considérerons dans la suite que des bases "orthonormées" dans ce sens. Notons également que le carré scalaire ne conduit pas à une norme définie positive (nous savons bien que le carré d'un intervalle de genre espace est négatif). Le tenseur métrique qui n'apparaît ici que comme une notation commode dans l'écriture du produit scalaire de deux événements est une des notions centrales en relativité générale, puisqu'il constitue la variable dynamique du champ de gravitation.

Cette écriture du produit scalaire peut être rendue plus compacte encore en introduisant la notion de coordonnée covariante.

[2]Remarquons à ce propos que le nom d'un indice muet n'a aucune importance dans l'écriture.

2.1.2 Coordonnées covariantes

Posons:

$$y_\mu = g_{\mu\nu} y^\nu \ . \tag{2.9}$$

Avant d'aller plus avant, commentons le fonctionnement des règles de sommation sur cette expression. L'indice répété (haut et bas) dans le membre de droite est ν. On doit donc sommer sur toutes les valeurs de cet indice. L'indice μ est un indice libre, qui apparaît sous le même nom et dans la même position (basse) dans les deux membres de l'équation. Nous appellerons "coordonnées covariantes de l'événement" les composantes y_μ.

En pratique, les règles de correspondance entre composantes contravariantes et composantes covariantes sont très simples: $y_0 = y^0$, $y_i = -y^i$. Les composantes covariantes nous permettent d'abaisser (ou d'élever pour la transformation inverse) les indices. Les écritures précédentes révèlent une règle très générale: l'abaissement ou l'élévation d'un indice spatial change le signe, alors que l'élévation ou l'abaissement d'un indice temporel ne s'accompagne pas d'un changement de signe.

Avec ces notations, le produit scalaire de deux 4–vecteurs s'écrit simplement:

$$\mathbf{R} \cdot \mathbf{S} = x^\mu y_\mu \tag{2.10}$$

mais aussi

$$\mathbf{R} \cdot \mathbf{S} = x_\mu y^\mu \qquad \text{avec} \quad x_\mu = g_{\mu\nu} x^\nu \ . \tag{2.11}$$

Le mérite de ces notations est donc de dissimuler les conventions de signes du tenseur métrique dans la définition des coordonnées covariantes et d'obtenir un produit scalaire s'écrivant "normalement". On peut bien sûr écrire la transformation inverse, donnant les coordonnées contravariantes en fonction des coordonnées covariantes. En définissant un nouveau "tableau de nombres" $g^{\mu\nu}$ par:

$$y^\nu = g^{\nu\mu} y_\mu \ , \tag{2.12}$$

on peut écrire:

$$y^\nu = g^{\nu\mu} y_\mu = g^{\nu\mu} g_{\mu\sigma} y^\sigma \ . \tag{2.13}$$

Les $g^{\mu\nu}$ sont donc tels que:

$$g^{\nu\mu} g_{\mu\sigma} = \delta^\nu_\sigma \ , \tag{2.14}$$

où les symboles δ^ν_σ sont de simples symboles de Kronecker (1 si les deux indices sont égaux, 0 sinon). En termes de matrices, la matrice des $g^{\mu\nu}$ est donc inverse de la matrice des $g_{\mu\nu}$. On a donc:

$$g^{\mu\nu} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} \ . \tag{2.15}$$

A titre d'exercice, montrons que les deux écritures du produit scalaire de l'équation (2.11) sont bien équivalentes:

$$x^\mu y_\mu = g^{\mu\nu} x_\nu g_{\mu\rho} y^\rho = g^{\mu\nu} g_{\mu\rho} x_\nu y^\rho = g^{\nu\mu} g_{\mu\rho} x_\nu y^\rho = \delta^\nu_\rho x_\nu y^\rho = x_\nu y^\nu \ , \tag{2.16}$$

la première et la dernière expression étant bien sûr équivalentes, puisque le nom d'un indice muet n'a aucune importance. Nous avons utilisé ici explicitement la symétrie du tenseur métrique par une permutation des indices.

2.1.3 Coordonnées covariantes, contravariantes et dualité

Nous avons introduit dans le paragraphe précédent les coordonnées covariantes comme une simple commodité de notation. En fait, elles ont une signification mathématique beaucoup plus profonde et plus fructueuse en termes de dualité. Nous allons donc rappeler quelques propriétés essentielles de la dualité (une notion d'algèbre linéaire qui joue aussi un rôle central en mécanique quantique).

Sur un espace vectoriel M, nous pouvons définir des *formes linéaires*. Une forme linéaire associe à tout vecteur un nombre réel (complexe dans le cas des espaces de Hilbert). On notera $\tilde{R}$ une forme linéaire et $\tilde{R}(\mathbf{S})$ le nombre réel associé au vecteur $\mathbf{S}$. Une forme linéaire est, comme son nom l'indique, une fonction linéaire de son argument vectoriel. On a donc des relations telles que: $\tilde{R}(\mathbf{S}+\mathbf{T}) = \tilde{R}(\mathbf{S}) + \tilde{R}(\mathbf{T})$ (le lecteur rétablira aisément l'ensemble des règles convenables).

On peut définir sur l'ensemble des formes linéaires une addition (simple addition des images) et la multiplication par un scalaire réel. Ces deux opérations confèrent à l'ensemble des formes linéaires une structure d'espace vectoriel. Nous l'appellerons l'espace *dual* de notre espace vectoriel initial (nous noterons M^* le dual de l'espace M).

On montre que, si M est de dimension finie, le dual est de même dimension. On peut de plus, si M est muni d'un produit scalaire, définir une bijection entre l'espace et le dual. On associe à chaque vecteur $\mathbf{R}$ la forme linéaire $\tilde{R}$ définie par $\tilde{R}(\mathbf{S}) = \mathbf{R}\cdot\mathbf{S}$. A chaque vecteur est associée une forme linéaire et chaque forme linéaire peut s'écrire comme un produit scalaire avec un vecteur fixe Les propriétés de linéarité du produit scalaire assurent que cette bijection est un isomorphisme entre l'espace et son dual. Cette relation très forte fait que l'on peut considérer un même objet soit comme un vecteur, soit comme une forme linéaire. En particulier, nous pourrons, dans l'espace temps à quatre dimensions, considérer au choix un événement comme un vecteur ou comme une forme linéaire. En fait, nous confondrons souvent les deux représentations en un seul objet.

Dans l'espace dual, nous pouvons choisir une base. En raison de l'isomorphisme avec l'espace d'origine, il est logique de prendre pour base les formes linéaires associées aux vecteurs de base de l'espace d'origine. En fait, pour des questions de notation, nous choisirons dans l'espace dual la base $\tilde{\epsilon}^\mu$ définie par:

$$\tilde{\epsilon}^\mu(\mathbf{e}_\nu) = \delta^\mu_\nu \ . \tag{2.17}$$

Notons que cette expression ne fait intervenir aucune sommation implicite. Notons également que les $\tilde{\epsilon}^\mu$ diffèrent des formes associées aux vecteurs de base par le signe, quand l'indice μ est spatial. Nous pouvons alors former, à partir d'un 4–vecteur événement $\mathbf{R} = x^\mu \mathbf{e}_\mu$, la forme linéaire $x_\mu \tilde{\epsilon}^\mu$. L'action de cette forme sur le 4–vecteur $\mathbf{S} = y^\mu \mathbf{e}_\mu$ s'écrit alors $x_\mu \tilde{\epsilon}^\mu(y^\nu \mathbf{e}_\nu) = x_\mu y^\mu = \mathbf{R}\cdot\mathbf{S}$. La forme ainsi construite coïncide donc avec la forme linéaire $\tilde{R}$ associée au vecteur $\mathbf{R}$. Si les composantes contravariantes sont les composantes du 4–vecteur, les composantes covariantes sont les composantes de la forme linéaire associée à ce vecteur sur la base duale. Ceci donne à ces composantes une signification mathématique beaucoup plus forte qu'une simple convention de signe. Comme nous pouvons confondre vecteur et forme linéaire en un seul objet physique, les composantes contravariantes et covariantes ne sont que deux écritures différentes d'une même quantité.

D'un point de vue "typographique", notons que les indices sont en bas pour les bases dans l'espace d'origine, en haut pour les bases dans l'espace dual, alors que les composantes dans l'espace d'origine (contravariantes) ont des indices en haut et les composantes dans l'espace dual (covariantes) des indices en bas. On comprend bien ici l'origine mathématique des règles de sommation sur les indices. N'écrire de sommations implicites que si l'un des indices est en haut et l'autre en bas, c'est reconnaître que la seule opération légitime est l'action d'une forme linéaire sur un vecteur (ou, autrement dit, le produit scalaire).

2.1.4 Changement de référentiel, changement de base

Un changement de référentiel transforme, comme nous l'avons vu au chapitre précédent, les coordonnées contravariantes d'un événement par la transformation de Lorentz. Nous allons d'abord ex-

primer cette transformation pour les coordonnées contravariantes et en déduire les transformations des coordonnées covariantes. Nous profiterons, pour ces écritures, de nos nouvelles notations. Nous reviendrons ensuite sur les transformations des vecteurs de base pour notre espace–temps et nous montrerons que la transformation de Lorentz est un simple changement de base orthonormée.

Les transformations que nous considérerons dans ce chapitre sont tout à fait générales. Elles englobent la forme spéciale de la transformation de Lorentz mais elles contiennent aussi les rotations, réflexions d'espace... constituant le groupe de Lorentz complet. Toutefois, quand nous expliciterons la forme des transformations, nous nous limiterons à la forme spéciale. Le passage au cas général ne pose que des problèmes d'écriture.

On peut écrire la transformation de Lorentz pour les coordonnées contravariantes comme:

$$x'^{\mu} = \mathcal{L}^{\mu}{}_{\nu} x^{\nu} \ , \tag{2.18}$$

où les x^{ν} sont les composantes de l'événement dans le référentiel $\mathcal{R}$ et les x'^{μ} les composantes contravariantes dans le nouveau référentiel $\mathcal{R}'$. En considérant, dans le tableau de nombres $\mathcal{L}^{\mu}{}_{\nu}$, l'indice μ (relatif donc au nouveau référentiel) comme un indice ligne et l'indice ν (relatif à l'ancien référentiel) comme un indice colonne, l'expression (2.18) apparaît comme un simple produit matriciel. Dans le cas de la forme spéciale de la transformation, on a simplement:

$$\mathcal{L}^{\mu}{}_{\nu} = \begin{pmatrix} \gamma & -\beta\gamma & 0 & 0 \\ -\beta\gamma & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \ , \tag{2.19}$$

qu'on pourrait encore écrire en termes de la rapidité ϕ, paramètre additif du groupe de Lorentz. La transformation inverse peut a priori s'écrire

$$x^{\mu} = (\mathcal{L}^{-1})^{\mu}{}_{\nu} x'^{\nu} \ , \ . \tag{2.20}$$

La matrice inverse $\mathcal{L}^{-1}$ s'obtenant simplement en changeant le signe de β dans les expressions précédentes. Nous verrons qu'en fait l'utilisation de cette matrice inverse n'est pas indispensable.

Examinons maintenant la loi de transformation des coordonnées covariantes (dans l'espace dual). Définissons un nouveau "tableau de nombres" $\mathcal{L}_{\mu}{}^{\nu}$ par:

$$x'_{\mu} = \mathcal{L}_{\mu}{}^{\nu} x_{\nu} \ . \tag{2.21}$$

Notons tout de suite que $\mathcal{L}_{\mu}{}^{\nu}$ est forcément différent de $\mathcal{L}^{\mu}{}_{\nu}$. L'ordre et la position (haute ou basse) des indices sont donc tous deux essentiels pour définir les quantités que nous manipulons. On peut déduire le lien entre $\mathcal{L}_{\mu}{}^{\nu}$ et $\mathcal{L}^{\mu}{}_{\nu}$ de l'invariance du produit scalaire (qui découle de l'invariance de l'intervalle). On a en effet:

$$x'^{\mu} y'_{\mu} = x^{\nu} y_{\nu} \qquad \text{avec} \quad x'^{\mu} = \mathcal{L}^{\mu}{}_{\rho} x^{\rho} \quad \text{et} \quad y'_{\mu} = \mathcal{L}_{\mu}{}^{\sigma} y_{\sigma} \ . \tag{2.22}$$

On en déduit donc:

$$\mathcal{L}^{\mu}{}_{\rho} x^{\rho} \mathcal{L}_{\mu}{}^{\sigma} y_{\sigma} = x^{\mu} y_{\mu} = x^{\rho} y_{\sigma} \delta^{\rho}_{\sigma} \ . \tag{2.23}$$

Cette relation devant être vérifiée pour tout couple de vecteurs, on a:

$$\mathcal{L}^{\mu}{}_{\rho} \mathcal{L}_{\mu}{}^{\sigma} = \delta^{\sigma}_{\rho} \ . \tag{2.24}$$

Notons que le membre de gauche de cette équation ne décrit pas un produit de matrices. En toute rigueur, on effectue la sommation sur deux indices lignes. Dans le cas de la forme spéciale de la transformation de Lorentz, les matrices sont symétriques et la matrice de transformation des coordonnées covariantes est simplement l'inverse de la matrice de transformation des coordonnées contravariantes,

inverse que l'on obtient en changeant le signe de la vitesse relative. On aurait pu établir ce résultat en remarquant simplement que le passage des coordonnées contravariantes aux coordonnées covariantes est un simple changement de signe pour les composantes spatiales. On change donc, dans la matrice de transformation, le signe des éléments spatio–temporels. Dans la forme spéciale de la transformation de Lorentz, cela revient à changer le signe de β.

On peut préciser encore le lien entre ces deux transformations en faisant intervenir le tenseur métrique dans l'expression du produit scalaire:

$$\mathbf{R} \cdot \mathbf{S} = x^\rho g_{\rho\sigma} y^\sigma = x'^\mu g_{\mu\nu} y'^\nu \ . \tag{2.25}$$

Le tenseur métrique qui exprime l'orthogonalité de la base est en effet évidemment invariant dans une transformation de Lorentz. En exprimant les nouvelles coordonnées en fonction des anciennes, on a:

$$x^\rho g_{\rho\sigma} y^\sigma = \mathcal{L}^\mu{}_\rho x^\rho g_{\mu\nu} \mathcal{L}^\nu{}_\sigma y^\sigma \ . \tag{2.26}$$

Cette relation étant vérifiée quelque soient $\mathbf{R}$ et $\mathbf{S}$, on en déduit:

$$g_{\rho\sigma} = \mathcal{L}^\mu{}_\rho g_{\mu\nu} \mathcal{L}^\nu{}_\sigma \ . \tag{2.27}$$

Le déterminant de la matrice représentant le tenseur métrique étant 1, on retrouve à partir de cette expression que le déterminant de la matrice représentant la transformation de Lorentz est ± 1 (rappelons qu'il n'est négatif que si la transformation fait intervenir une réflexion d'espace ou de temps). On peut mettre la dernière relation sous une forme plus parlante en multipliant les deux membres par $g^{\tau\rho}$ et en effectuant les sommations:

$$g^{\tau\rho} g_{\rho\sigma} = \delta^\tau_\sigma = g^{\tau\rho} \mathcal{L}^\mu{}_\rho g_{\mu\nu} \mathcal{L}^\nu{}_\sigma \ , \tag{2.28}$$

et donc:

$$\left(g^{\tau\rho} g_{\mu\nu} \mathcal{L}^\mu{}_\rho\right) \mathcal{L}^\nu{}_\sigma = \delta^\tau_\sigma \ . \tag{2.29}$$

En rapprochant cette équation de (2.24) en utilisant la symétrie de g, on voit que:

$$\mathcal{L}_\nu{}^\tau = g_{\nu\mu} g^{\tau\rho} \mathcal{L}^\mu{}_\rho \ . \tag{2.30}$$

D'un point de vue purement typographique, l'interprétation de cette relation est transparente. Nous savons, pour les coordonnées simples, que l'application d'un tenseur métrique avec les indices hauts transforme une coordonnée covariante (indice bas) en un coordonnée contravariante (indice haut). Elever les indices nécessite l'action d'un tenseur métrique à indices hauts (nous dirons complètement contravariant). De même un tenseur métrique complètement covariant (indices bas) abaisse l'indice, transformant une coordonnée contravariante en une coordonnée covariante. La relation que nous venons d'établir pour les changements de base généralise ce principe. Pour passer de $\mathcal{L}^\mu{}_\rho$, avec le premier indice en haut et le second en bas, à $\mathcal{L}_\nu{}^\tau$, il faut appliquer deux tenseurs métriques, l'un complètement contravariant qui élève l'indice colonne, l'autre complètement covariant qui abaisse l'indice ligne. Si on fait confiance à ces règles typographiques, l'établissement détaillé de (2.30) est inutile et on peut écrire cette transformation a priori. On écrira ainsi la relation inverse:

$$\mathcal{L}^\nu{}_\tau = g_{\tau\rho} g^{\nu\mu} \mathcal{L}_\mu{}^\rho \ . \tag{2.31}$$

Remarquons que la forme très simple du tenseur métrique rend triviales les modifications dans un abaissement ou une élévation d'indice. Comme pour les coordonnées, le changement de position d'un indice spatial change le signe alors qu'un indice temporel ne change pas le signe. Dans le passage d'une transformation à l'autre, seuls changent de signe les coefficients spatio–temporels. Les coefficients spatiaux et temporels sont inchangés. C'est bien ce qu'on observe sur la forme spéciale de la transformation de Lorentz.

On pourrait penser alors ne faire agir sur les transformations de Lorentz qu'un tenseur métrique et à définir deux nouvelles quantités:

$$\mathcal{L}^{\mu\tau} = g^{\tau\rho}\mathcal{L}^{\mu}{}_{\rho} \tag{2.32}$$
$$\mathcal{L}_{\mu\sigma} = g_{\mu\nu}\mathcal{L}^{\nu}{}_{\sigma} \ . \tag{2.33}$$

En fait ces nouvelles quantités décrivent, comme nous allons le voir, le passage simultané d'un référentiel à l'autre et d'un type de coordonnées à l'autre. On a en effet:

$$x'^{\mu} = \mathcal{L}^{\mu}{}_{\rho}x^{\rho} = \mathcal{L}^{\mu}{}_{\rho}g^{\rho\nu}x_{\nu} = \mathcal{L}^{\mu\nu}x_{\nu} \tag{2.34}$$

et de même:

$$x'_{\mu} = \mathcal{L}_{\mu\nu}x^{\nu} \ . \tag{2.35}$$

On peut enfin utiliser la relation (2.24) pour exprimer le changement de référentiel inverse sans faire intervenir la transformation $\mathcal{L}^{-1}$. On a en effet:

$$\mathcal{L}^{\mu}{}_{\rho}x'_{\mu} = \mathcal{L}^{\mu}{}_{\rho}\mathcal{L}_{\mu}{}^{\nu}x_{\nu} = \delta^{\nu}_{\rho}x_{\nu} = x_{\rho} \tag{2.36}$$

et donc

$$x_{\rho} = \mathcal{L}^{\mu}{}_{\rho}x'_{\mu} \ . \tag{2.37}$$

On peut bien sûr écrire des transformations similaires pour les coordonnées contravariantes ou toute combinaison de composantes mixtes:

$$x'_{\mu} = \mathcal{L}_{\mu}{}^{\nu}x_{\nu} \tag{2.38}$$
$$x'^{\mu} = \mathcal{L}^{\mu}{}_{\nu}x^{\nu} \tag{2.39}$$
$$x_{\mu} = \mathcal{L}^{\nu}{}_{\mu}x'_{\nu} \tag{2.40}$$
$$x^{\mu} = \mathcal{L}_{\nu}{}^{\mu}x'^{\nu} \ , \tag{2.41}$$

les différentes expressions de la transformation se déduisant de la forme originale par les règles d'élévation et d'abaissement des indices. Ces différentes combinaisons s'obtiennent automatiquement (ou plutôt typographiquement) en respectant les règles de balancement des indices, en attribuant le premier indice (ligne) au nouveau référentiel, le second à l'ancien et en sommant sur l'indice correspondant au référentiel de la coordonnée à transformer. L'ensemble de ces règles très strictes, que nous avons justifiées en détail, minimise le risque d'erreurs dans ces écritures, beaucoup plus efficacement qu'avec les notations standard de l'algèbre linéaire.

Nous terminerons ce paragraphe en examinant les lois de transformation des vecteurs de base de notre espace–temps. De manière évidente, la transformation de Lorentz correspond à un simple changement de base dans l'espace temps. Le même événement, le même vecteur, s'exprime par deux jeux de composantes différentes quand on le représente dans deux référentiels ou deux bases différentes. En remarquant que $x_{\nu} = \mathbf{R}\cdot\mathbf{e}_{\nu} = \tilde{R}(\mathbf{e}_{\nu})$ et que $x'_{\mu} = \mathbf{R}\cdot\mathbf{e}'_{\mu}$, où les $\mathbf{e}'_{\mu}$ sont les transformés des vecteurs de base, on peut écrire $\mathbf{R}\cdot\mathbf{e}'_{\mu} = \mathcal{L}_{\mu}{}^{\nu}\ \mathbf{R}\cdot\mathbf{e}_{\nu}$, relation qui doit être vérifiée pour tout vecteur $\mathbf{R}$. On en déduit donc que:

$$\mathbf{e}'_{\mu} = \mathcal{L}_{\mu}{}^{\nu}\mathbf{e}_{\nu} \ . \tag{2.42}$$

En d'autres termes, la loi de transformation des vecteurs de base est celle des composantes covariantes, inverse (au sens précisé plus haut) de la transformation des composantes contravariantes. Les appellations covariantes et contravariantes proviennent précisément de ces comportements.

On peut aussi exprimer la loi de transformation de la base duale. En écrivant $x'^{\mu} = \tilde{\epsilon}'^{\mu}(\mathbf{R})$, $x^{\mu} = \tilde{\epsilon}^{\mu}(\mathbf{R})$, on en déduit:

$$\tilde{\epsilon}'^{\mu} = \mathcal{L}^{\mu}{}_{\nu}\tilde{\epsilon}^{\nu} \ . \tag{2.43}$$

Les vecteurs de la base duale se transforment donc comme les composantes contravariantes.

2.2 Autres 4–vecteurs

2.2.1 Définition

Nous n'avons considéré jusque là que le 4–vecteur associé à la position d'un événement dans l'espace–temps. Nous allons maintenant généraliser la notion à d'autres quantités physiques. Nous appellerons 4–vecteur toute quantité physique s'exprimant par 4 coordonnées (covariantes ou contravariantes) se transformant comme le 4–vecteur position dans un changement de référentiel (c'est à dire par une transformation de Lorentz). De manière évidente, toute collection de quatre quantités physiques arbitraires ne constitue pas un 4–vecteur. Nous verrons dans les prochains paragraphes quelques exemples de 4–vecteurs.

Les composantes contravariantes de tout 4–vecteur $\mathbf{A}$ peuvent s'écrire sous la forme de trois composantes spatiales qui forment un vecteur (ce qui résulte des lois de transformation dans le sous groupe du groupe de Lorentz qui décrit les changements de repère sans changement de référentiel) et d'une coordonnée temporelle. Nous noterons $\mathbf{A} = (a^0, \mathbf{a})$, où a^0 est sa composante temporelle et $\mathbf{a}$ le vecteur tridimensionnel. Les composantes covariantes pourraient alors se mettre sous la forme $(a_0 = a^0, -\mathbf{a})$. A priori, ces composantes ont une dimension qui peut être différente de celles des composantes du 4–vecteur position. Chaque espèce de 4–vecteur devrait donc, en toute rigueur, exister dans un espace–temps propre. En fait, comme en dynamique classique, nous ferons évoluer tous nos 4–vecteurs, positions, vitesses et accélérations dans le même espace. Nous prendrons simplement garde à ne pas additionner des objets de natures différentes.

Les notions de produit scalaire, de composantes covariantes et contravariantes, les règles de changement de base ou de référentiel s'appliquent bien sûr aux 4–vecteurs quelle que soit leur nature.

Nous profiterons aussi de ce paragraphe pour introduire la notion de 4–scalaire. Il s'agit simplement d'une quantité physique indépendante du référentiel, telle que la vitesse limite, la constante de structure fine, la masse d'une particule ou, plus simplement, le produit scalaire de deux 4–vecteurs (le produit scalaire de vecteurs de dimensions différentes n'étant pas exclu). Le produit d'un 4–vecteur par un 4–scalaire donne bien sûr un autre 4–vecteur.

2.2.2 4–vitesse, 4–impulsion, 4–accélération

Nous chercherons ici à définir la vitesse d'une particule relativiste. Il nous faut pour cela paramétrer la ligne d'univers de la particule par un paramètre temporel. Ce paramètre doit être un 4–scalaire. Nous pourrons ainsi définir simplement la vitesse comme étant le rapport de l'accroissement du 4–vecteur position pendant un petit intervalle de temps à l'accroissement de ce paramètre temporel. Le résultat sera alors évidemment un 4–vecteur. Le temps absolu d'un référentiel ne convient pas pour paramétrer la trajectoire. Nous avons vu, en revanche, que le temps propre τ de la particule en mouvement est indépendant du choix de référentiel. Ce temps propre permet donc de définir un 4–vecteur vitesse par:

$$\mathbf{U} = \frac{d\mathbf{R}}{d\tau} \ . \tag{2.44}$$

Ecrivons cette quantité en fonction de la vitesse spatiale ordinaire $\mathbf{v}$ de la particule dans un référentiel $\mathcal{R}$ donné. Si le temps propre varie de $d\tau$, le temps dans $\mathcal{R}$ varie de $dt = \gamma d\tau$ (dilatation des temps). La position de la particule variant de $d\mathbf{r} = \mathbf{v}dt$, on a immédiatement $d\mathbf{R}/dt = (c, \mathbf{v})$ et

$$\mathbf{U} = (c\gamma, \gamma\mathbf{v}) \ . \tag{2.45}$$

Remarquons que la partie spatiale de la 4–vitesse n'est autre que la célérité (vitesse calculée dans l'espace du référentiel et dans le temps propre de la particule). $\mathbf{U}$ étant un 4–vecteur, il se transforme par la transformation de Lorentz dans un changement de référentiel. Un excellent exercice consiste à appliquer les lois de transformations décrites au paragraphe précédent pour retrouver la loi de

composition des vitesses relativistes. Notons enfin que le carré de la norme de $\mathbf{U}$, $U^\mu U_\mu = c^2\gamma^2 - \gamma^2 v^2 = c^2$ n'est autre que le carré de la vitesse de la lumière.

En multipliant la 4–vitesse par la masse de la particule, qui est évidemment un 4–scalaire, on obtient encore un 4–vecteur, l'impulsion de la particule:

$$\mathbf{P} = m\mathbf{U} = (p^0, \mathbf{p}) = (m\gamma c, m\gamma \mathbf{v}) \ . \tag{2.46}$$

Cette quantité jouera un rôle essentiel dans notre nouvelle dynamique.

La 4–vitesse $\mathbf{U}$ est en général une fonction du temps propre τ de la particule. Nous pouvons donc la dériver encore par rapport à ce temps, pour aboutir à une définition de la 4–accélération

$$\mathbf{\Gamma} = \frac{d\mathbf{U}}{d\tau} = \gamma \frac{d\mathbf{U}}{dt} = (c\gamma\gamma', \gamma\gamma'\mathbf{v} + \gamma^2\mathbf{a}) \ , \tag{2.47}$$

où $\gamma' = d\gamma/dt$ est la dérivée temporelle ordinaire du facteur γ et $\mathbf{a}$ est l'accélération tridimensionnelle de la particule. On remarque que la 4–accélération est perpendiculaire à la 4–vitesse: $\Gamma^\mu U_\mu = 0$. Si on peut établir cette relation à partir des expressions précédentes de ces 4–vecteurs, elle découle beaucoup plus simplement du fait que le module de $\mathbf{U}$ est une constante.

2.2.3 Densité de courant

Nous allons maintenant former un 4–vecteur à partir de la densité de charges ρ et de la densité de courant $\mathbf{j}$. Il est clair en effet que ces deux quantités se transforment de manières profondément reliées dans un changement de référentiel. Pour comprendre comment former ce 4–vecteur, nous allons considérer le cas d'une charge q "ponctuelle", que nous modéliserons comme une densité de charge constante contenue dans un petit élément de volume dV (défini dans un référentiel $\mathcal{R}$). Nous écrirons donc $q = \rho dV$. La particule se déplace à la vitesse $\mathbf{v}$ dans $\mathcal{R}$. La densité de courant est donc $\mathbf{j} = \rho\mathbf{v}$ en tous points de dV (elle est nulle en dehors).

Dans un autre référentiel $\mathcal{R}'$, toutes ces quantités sont a priori modifiées. La "particule" occupe, à un instant donné dans $\mathcal{R}'$, un élément de volume dV'. Seule la charge totale q, qui est une propriété intrinsèque de la particule, est un 4–scalaire. On a donc $\rho dV = \rho' dV'$. En d'autres termes, ρdV est un 4–scalaire (notons que le lien entre dV et dV' n'est pas évident, puisqu'il faut faire la transformation de Lorentz sur tous les "coins" de l'élément de volume. Nous reviendrons bientôt sur ce point). Considérons maintenant un intervalle de temps dt dans $\mathcal{R}$. Pendant ce temps, la particule se déplace de $dx^\mu = (cdt, d\mathbf{r})$ qui est évidemment un 4–vecteur. La quantité $\rho dV dx^\mu = \rho dt dV (dx^\mu/dt) = \rho(d\Omega/c)dx^\mu/dt$ est donc, elle aussi un 4–vecteur, si $d\Omega = cdtdV$ représente un élément d'intégration dans l'espace–temps (en revanche, dx^μ/dt n'est pas un 4–vecteur, de même que dt ou dV pris séparément ne sont pas des 4–scalaires).

Nous montrerons maintenant que $d\Omega = cdtdV$ est un 4–scalaire. D'un point de vue physique, $d\Omega$ représente un petit domaine de l'espace–temps considéré entre deux instants infiniment voisins et a la dimension d'un volume multiplié par un temps. Modélisons ce petit domaine par un cube élémentaire de dimensions dx, dy et dz (voir Fig. 2.1). Si nous choisissons les axes convenablement, nous n'aurons à considérer que la forme spéciale de la transformation de Lorentz. Les coordonnées y et z étant inchangées, nous nous ramenons à démontrer l'invariance de l'élément de surface dans le plan (x, t). L'étendue de notre élément est la surface rectangulaire de côtés dx et dt dans le référentiel $\mathcal{R}$. Sa mesure est bien $dxdt$. Remarquons tout de suite que $dxdt$ n'est pas un invariant de Lorentz, comme on peut s'en convaincre aisément. Vu dans le référentiel $\mathcal{R}'$, le petit élément de surface prend l'aspect représenté sur la partie droite de la figure 2.1. Il s'agit d'un petit losange. La surface de ce losange peut s'exprimer facilement en écrivant les transformés de Lorentz de trois des sommets. En fait, on trouve que cette mesure, qui est manifestement différente de $dx'dt'$ est égale à $dxdt$ (ce résultat découle directement du fait que le déterminant de la transformation de Lorentz est un, ce qui confirme sa généralité). L'étendue totale dans l'espace–temps d'un petit élément d'intégration est

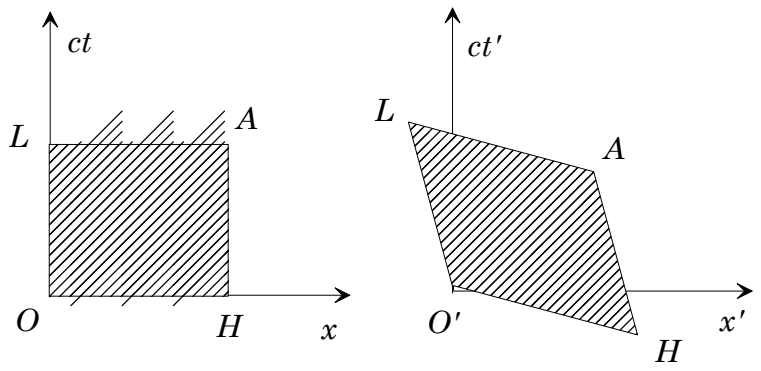

Figure 2.1: Un même élément de volume de l'espace–temps vu dans deux référentiels. La forme change, mais la surface (ou l'hypervolume dans l'espace quadridimensionnel) est constante.

donc un 4–scalaire (résultat qui nous permettra plus tard de définir convenablement des intégrales de volume dans l'espace–temps).

On déduit de tout ce qui précède que

$$\mathbf{J}^\mu = \rho \frac{dx^\mu}{dt} \tag{2.48}$$

est un 4–vecteur réunissant les densités de charge et de courant créés par une charge ponctuelle. On peut aussi écrire

$$\mathbf{J}^\mu = (c\rho, \mathbf{j}) \ , \tag{2.49}$$

en faisant intervenir la densité de courant tridimensionnelle. Pour une répartition quelconque de charges et de courants, résultant de la superposition des mouvements d'un grand nombre de particules ponctuelles, la quantité écrite en termes des densités totales de charges et de courants est bien sûr encore un 4–vecteur.

2.2.4 Vecteur d'onde

Considérons, dans le référentiel $\mathcal{R}$, une onde plane monochromatique de fréquence ω et de vecteur d'onde $\mathbf{k}$. Les différents champs et potentiels de cette onde sont tous proportionnels à $\exp(-i\phi)$ avec:

$$\phi = \omega t - \mathbf{k} \cdot \mathbf{r} \ . \tag{2.50}$$

La phase ϕ de l'onde plane se doit d'être un 4–scalaire. Sa valeur pour un événement donné représente en effet la valeur relative commune des champs et potentiels par rapport à leur maximum. Les transformations de Lorentz de ces quantités, dont nous ne préjugerons pas encore, étant sûrement linéaires, le même facteur $\exp(i\phi)$ décrit, dans un autre référentiel $\mathcal{R}'$, la valeur des quantités transformées par rapport à leurs nouvelles amplitudes.

En posant

$$k^\mu = (\omega/c, \mathbf{k}) \ , \tag{2.51}$$

on peut, au moins formellement, écire:

$$\phi = k_\mu x^\mu \ , \tag{2.52}$$

où les x^μ sont les composantes contravariantes de l'événement considéré. ϕ étant un 4–scalaire et x^μ un 4–vecteur, il est évident que k^μ est, lui aussi, un 4–vecteur généralisant le vecteur d'onde dans l'espace–temps. On notera que sa norme est nulle:

$$k_\mu k^\mu = 0 \tag{2.53}$$

puisque $\omega^2 - c^2k^2 = 0$ pour une onde plane dans le vide. Le vecteur d'onde se transforme donc simplement par une transformation de Lorentz quand on passe d'un référentiel à un autre:

$$k'^{\mu} = \mathcal{L}^{\mu}{}_{\nu} k^{\nu} \ . \tag{2.54}$$

En explicitant cette transformation pour les coordonnées spatiales et temporelle, on obtient:

$$\omega' = \gamma(\omega - uk_x) \tag{2.55}$$
$$k'_x = \gamma(k_x - \beta\omega/c) \tag{2.56}$$
$$k'_y = k_y \tag{2.57}$$
$$k'_z = k_z \tag{2.58}$$

Dans la première équation, on reconnaît le changement de fréquence d'une onde dans un changement de référentiel, c'est à dire l'effet Doppler. L'expression que nous trouvons ici ne diffère de l'expression habituelle en relativité galiléenne ($\omega' = \omega - \mathbf{k} \cdot \mathbf{u}$) que par le facteur de dilatation des temps γ (il est bien naturel que ce facteur intervienne dans la modification de la fréquence). Pour des mobiles qui ne sont pas trop rapides (ce qui est presque toujours le cas, par exemple, pour des sources atomiques), l'effet Doppler reste dominé par l'effet classique. Il est un cas cependant où l'effet relativiste domine, celui où la vitesse $\mathbf{u}$ est strictement perpendiculaire au vecteur d'onde. En cinématique classique, la fréquence n'est pas affectée. En relativité einsteinienne, en revanche, la fréquence est multipliée par le facteur γ de dilatation des temps. La modification relative de la fréquence est donc du second ordre en β (d'où le nom d'effet Doppler du second ordre), alors que l'effet Doppler classique est du premier ordre. C'est donc un effet faible pour des vitesses ordinaires qui est cependant gênant dans des expériences de haute précision. La spectroscopie atomique la plus précise actuellement est réalisée sur des atomes d'hydrogène excités par laser. Pour s'affranchir de l'effet Doppler, on fait en sorte que la vitesse des atomes soit très précisément perpendiculaire à la direction du laser. La cause essentielle d'erreur devient alors l'effet Doppler du second ordre qu'il faut corriger très précisément. Pour cela, il est indispensable de réduire autant que possible la vitesse des atomes et de la déterminer précisément.

Les équations portant sur la partie spatiale décrivent le changement de direction de propagation. Les composantes transverses ne sont pas modifiées. La composante longitudinale change de telle manière que la relation de dispersion des ondes planes dans le vide $k' = \omega'/c$ reste vérifiée, comme on pourra s'en convaincre aisément. Notons qu'on peut ainsi retrouver de manière très rapide l'aberration relativiste des étoiles. La situation que nous considérions au chapitre précédent était en effet une onde se propageant vers l'origine le long de l'axe Oy: k_y est donc la seule composante non nulle dans le référentiel $\mathcal{R}$. En appliquant la transformation précédente, on trouve immédiatement:

$$k'_y = k_y = -\omega/c \tag{2.59}$$
$$k'_x = -\beta\gamma\omega/c \tag{2.60}$$

La direction de propagation fait donc, dans le référentiel $\mathcal{R}'$ un angle $\theta = \arctan \beta\gamma$ avec l'axe Oy'. Nous retrouvons bien le résultat du chapitre précédent.

2.3 Tenseurs

Nous aurons à manipuler des quantités plus complexes que les 4–vecteurs. Dans un espace vectoriel, on peut bien sûr définir des opérateurs linéaires, représentés par des matrices dans une base convenable. En fait, nous considérerons des objets plus généraux pouvant dépendre de plus de deux indices. De tels objets décrivent des lois de composition multilinéaires entre vecteurs ou opérateurs. Les tenseurs se retrouvent en fait dans de nombreux domaines de la physique. En élasticité, par exemple, la déformation d'un solide dans une direction peut dépendre d'une contrainte appliquée dans une autre

direction. La relation linéaire correspondante est décrite par un tenseur de rang 2. En optique non linéaire, un ensemble de trois champs électriques peut créer une polarisation dans le milieu. Il faudra donc écrire une application linéaire donnant un vecteur à partir de trois autres, ce qui doit être décrit par un tenseur à quatre indices. Notons tout de suite que n'importe quel tableau de nombres n'est pas un tenseur, comme toute collection de nombres ne représente pas forcément un vecteur. Il faut, de plus, que ces quantités se transforment "bien" dans un changement de base.

2.3.1 Tenseurs contravariants

L'opération de produit tensoriel permet d'associer à l'espace vectoriel M un espace $M \otimes M$, plus grand. A tout couple de vecteurs $\mathbf{R}$ et $\mathbf{S}$ de M, on associe un vecteur $\mathbf{R} \otimes \mathbf{S}$ de $M \otimes M$. Une base de $M \otimes M$ est formée des 16 produits tensoriels possibles formés avec les 4 vecteurs de base de M: $\mathbf{e}_\mu \otimes \mathbf{e}_\nu$. Les composantes de $\mathbf{R} \otimes \mathbf{S}$ sur cette base sont les produits des composantes de $\mathbf{R}$ et $\mathbf{S}$:

$$\mathbf{R} \otimes \mathbf{S} = x^\mu y^\nu \mathbf{e}_\mu \otimes \mathbf{e}_\nu \ . \tag{2.61}$$

La dimension de l'espace produit tensoriel est 16.Un produit tensoriel de deux vecteurs ne dépend que de 8 paramètres libres. Nous pouvons donc définir des objets plus généraux que les produits tensoriels de vecteurs: les tenseurs de rang 2 complètement contravariants, éléments de l'espace $M \otimes M$. Nous noterons $T^{\mu\nu}$ les composantes d'un tel tenseur sur la base $\mathbf{e}_\mu \otimes \mathbf{e}_\nu$. En fait, nous avons déjà rencontré de tels objets, par exemple avec le tenseur métrique complètement contravariant.

On peut déduire la règle de transformation d'un tenseur dans un changement de base (c'est à dire une transformation de Lorentz) de celle d'un produit tensoriel de vecteurs: $T^{\mu\nu}$ se transforme comme un produit de composantes contravariantes:

$$T'^{\mu\nu} = \mathcal{L}^\mu{}_\rho \mathcal{L}^\nu{}_\sigma T^{\rho\sigma} \ . \tag{2.62}$$

En appliquant les règles du paragraphe précédent, nous aurons aussi la transformation inverse:

$$T^{\mu\nu} = \mathcal{L}_\rho{}^\mu \mathcal{L}_\sigma{}^\nu T'^{\rho\sigma} \ . \tag{2.63}$$

Notons que ces expressions ne décrivent pas des produits de trois matrices au sens ordinaires. Dans la première, par exemple, la sommation sur σ est une sommation sur un indice colonne dans les deux termes où il apparaît. Il faut donc prendre garde, dans les calculs pratiques, d'effectuer les transpositions nécessaires si on veut utiliser les règles standard du produit matriciel.

L'opération de produit tensoriel peut être généralisée aisément à un nombre arbitraires de termes. On peut définir l'espace $M^{\otimes k}$, produit tensoriel de M k fois avec lui même. Les éléments de cet espace, de dimension 4^k, sont les tenseurs complètement contravariants de rang k et leurs composantes s'écriront $T^{\mu\nu\rho\cdots\tau}$. Ces composantes se transforment par un "produit" de k transformations de Lorentz. Nous laissons au lecteur le soin de l'écrire. Un tenseur de rang n décrit une transformation multilinéaire qui, à $n-1$ vecteurs associe un vecteur.

2.3.2 Tenseurs covariants, tenseurs mixtes

Ce que nous avons fait pour l'espace M peut être repris pour son dual M^*. On peut définir ainsi des tenseurs de rang deux, complètement covariants, dont les composantes sur la base produit tensoriel $\tilde{\epsilon}^\mu \otimes \tilde{\epsilon}^\nu$ s'écriront $T_{\mu\nu}$. La transformation de Lorentz de ces quantités s'écrit simplement:

$$T'_{\mu\nu} = \mathcal{L}_\mu{}^\rho \mathcal{L}_\nu{}^\sigma T_{\rho\sigma} \ . \tag{2.64}$$

On peut faire le produit tensoriel d'un nombre arbitraires d'espaces duaux. On peut aussi définir des objets appartenant au produit tensoriel de l'espace M par son dual M^*. On obtient alors des tenseurs mixtes de rang deux (ou plus si on utilise plusieurs M et M^*) dont les composantes s'écriront

$T^{\mu}{}_{\nu}$ (pour $M \otimes M^*$) ou $T_{\mu}{}^{\nu}$ (pour $M^* \otimes M$). Notons que ces deux écritures recouvrent a priori des objets différents. La règle de transformation d'un tel tenseur mixte est simplement:

$$T'^{\mu}{}_{\nu} = \mathcal{L}^{\mu}{}_{\rho}\mathcal{L}_{\nu}{}^{\sigma}T^{\rho}{}_{\sigma} \ , \tag{2.65}$$

et se généralise aisément à tout tenseur mixte de n'importe quel rang.

En fait, nous savons bien que les composantes contravariantes et les composantes covariantes recouvrent le même objet physique. Il en est de même pour les tenseurs: une quantité physique s'exprimant comme un tenseur peut être écrite à volonté comme un tenseur complètement contravariant, complètement covariant, ou mixte de façon arbitraire. Comme pour les composantes des 4–vecteurs (et cela résulte du lien entre tenseur et produits de ces composantes), le tenseur métrique $g^{\mu\nu}$ ou $g_{\mu\nu}$ peut être utilisé pour élever ou abaisser les indices. Nous pourrons écrire par exemple:

$$T^{\mu\nu} = g^{\mu\rho}g^{\nu\sigma}T_{\rho\sigma} \tag{2.66}$$

$$T^{\mu}{}_{\nu} = g^{\mu\rho}g_{\nu\sigma}T_{\rho}{}^{\sigma} \tag{2.67}$$

$$T^{\mu}{}_{\nu} = g_{\nu\rho}T^{\mu\rho} \ . \tag{2.68}$$

Le tenseur métrique se réduisant en relativité restreinte, à un tableau diagonal de signes, ces règles de transformation prennent une forme très simple: l'abaissement ou l'élévation d'un indice spatial change le signe de la quantité, alors que l'élévation ou l'abaissement d'un indice temporel ne change rien. On trouve ainsi, par exemple, que seuls les indices spatio–temporels changent de signe dans le passage d'un tenseur de rang 2 de la forme complètement contravariante à la forme complètement covariante.

Ces définition des tenseurs peut paraître abstraites. On peut aussi voir les tenseurs de rang deux comme des applications linéaires de l'espace M dans lui même (ou de son dual dans lui même). Un tenseur complètement covariant de rang 2, par exemple, décrit naturellement une application linéaire de M dans M^*. L'image W d'un vecteur V s'écrit alors:

$$W_{\mu} = T_{\mu\nu}V^{\nu} \ . \tag{2.69}$$

On aurait des écritures similaires pour les tenseurs complètement contravariants (de M^* dans M) ou les tenseurs mixtes, décrivant la transformation entre deux objets de même nature. Nous laisserons au lecteur le soin de les écrire.

Les règles de changement de base pour les tenseurs peuvent se déduire de cette forme. Nous prendrons une transformation de Lorentz $\mathcal{L}$. Les transformations des composantes de W et de V s'écrivent:

$$W'_{\rho} = \mathcal{L}_{\rho}{}^{\mu}W_{\mu} \qquad\qquad V^{\nu} = \mathcal{L}_{\sigma}{}^{\nu}V'^{\sigma} \ . \tag{2.70}$$

On peut donc écrire:

$$W'_{\rho} = \mathcal{L}_{\rho}{}^{\mu}W_{\mu} = \mathcal{L}_{\rho}{}^{\mu}T_{\mu\nu}V^{\nu} = \mathcal{L}_{\rho}{}^{\mu}T_{\mu\nu}\mathcal{L}_{\sigma}{}^{\nu}V'^{\sigma} = T'_{\rho\sigma}V'^{\sigma} \ . \tag{2.71}$$

On lit directement sur cette équation l'expression des composantes transformées du tenseur. On constatera qu'elles coïncident bien avec celles que nous avons obtenues plus haut.

2.3.3 Vocabulaire et exemples

Un tenseur de rang 2 est dit symétrique si:

$$T^{\mu\nu} = T^{\nu\mu} \tag{2.72}$$

On en déduit immédiatement $T_{\mu\nu} = T_{\nu\mu}$ et $T^{\mu}{}_{\nu} = T_{\nu}{}^{\mu}$. On pourra donc écrire la forme mixte sans préciser l'ordre des indices, comme T^{μ}_{ν}. Notons que la symétrie du tenseur n'implique *pas* l'égalité de T^{μ}_{ν} et de T^{ν}_{μ} (il n'est que d'examiner le cas des indices spatio–temporels pour s'en convaincre).

Un tenseur de rang 2 est dit antisymétrique si:

$$T^{\mu\nu} = -T^{\nu\mu} \ . \tag{2.73}$$

Les termes diagonaux de ce tenseur $T^{\mu\mu}$ sont évidemment nuls. On montre qu'un tenseur antisymétrique peut se mettre sous la forme:

$$T^{\mu\nu} = \begin{pmatrix} 0 & a_x & a_y & a_z \\ -a_x & 0 & -b_z & b_y \\ -a_y & b_z & 0 & -b_x \\ -a_z & -b_y & b_x & 0 \end{pmatrix} = (\mathbf{a}, \mathbf{b}) \tag{2.74}$$

où $\mathbf{a}$ est un vecteur spatial et $\mathbf{b}$ un pseudo-vecteur (qui se transforme en l'opposé de son symétrique dans un changement de base comprenant une réflexion d'espace). Le couple champ électrique/champ magnétique obéissant à ces conditions, nous ne nous étonnerons pas que le champ électromagnétique s'exprime comme un tenseur antisymétrique de rang 2.

Nous appellerons "trace" d'un tenseur de rang 2 la quantité $T^{\mu}{}_{\mu} = T_{\mu}{}^{\mu}$. Plus généralement, nous appellerons "contraction" d'un tenseur sur un indice une expression comme $T^{\mu}{}_{\mu}{}^{\rho}$. La contraction d'un tenseur de rang k sur un indice est un tenseur de rang $k-2$ (la trace étant une contraction d'un tenseur de rang 2, elle donne un tenseur de rang 0, c'est à dire un 4–scalaire). La contraction d'un tenseur de rang trois donne, pour sa part, un tenseur de rang 1, c'est à dire un 4–vecteur.

Montrons, à titre d'exercice, que la contraction sur un indice d'un tenseur de rang trois se transforme bien comme un vecteur:

$$\begin{aligned} T'^{\mu}{}_{\mu}{}^{\rho} &= \mathcal{L}^{\mu}{}_{\sigma}\mathcal{L}_{\mu}{}^{\nu}\mathcal{L}^{\rho}{}_{\tau}T^{\sigma}{}_{\nu}{}^{\tau} \\ &= \delta^{\nu}_{\sigma}\mathcal{L}^{\rho}{}_{\tau}T^{\sigma}{}_{\nu}{}^{\tau} \\ &= \mathcal{L}^{\rho}{}_{\tau}T^{\sigma}{}_{\sigma}{}^{\tau} \ , \end{aligned} \tag{2.75}$$

ce qui établit bien la propriété cherchée.

Comme tenseurs, nous connaissons déjà le tenseur métrique. On pourra vérifier directement qu'il est bien invariant dans une transformation de Lorentz. Il s'agit d'un tenseur symétrique. Sa forme mixte est évidemment: $g^{\mu}_{\nu} = g^{\mu\rho}g_{\nu\rho} = \delta^{\mu}_{\nu}$. Le symbole de Kronecker n'est donc que la forme mixte du tenseur métrique! La relation entre les formes contravariantes et covariantes $g^{\mu\nu}g_{\nu\rho} = \delta^{\mu}_{\rho}$ peut donc s'interpréter comme un simple abaissement d'indice.

Notons à ce point que les opérateurs $\mathcal{L}$ de changement de base ont toutes les caractéristiques de tenseurs et se comportent normalement vis à vis de l'élévation ou de l'abaissement des indices. En toute rigueur, nous ne devons pas les considérer comme des tenseurs. On sait bien, en algèbre linéaire, que les matrices de changement de base ne sont pas à proprement parler des opérateurs. On peut aussi s'en convaincre en s'interrogeant sur le sens physique d'un changement de base pour une transformation de Lorentz.

Nous définirons finalement le tenseur de rang 4 complètement antisymétrique $\epsilon^{\mu\nu\rho\sigma}$. Parmi les 256 éléments de ce tenseur, seuls sont non nuls ceux dont les indices correspondent à une permutation de $(0, 1, 2, 3)$. Si la permutation est paire, l'élément correspondant vaut $+1$. Il vaut -1 si la permutation est impaire. Il n'y a donc que 24 éléments non nuls dans ϵ, 12 valant $+1$ et 12 -1. On a $\epsilon^{\mu\nu\rho\sigma} = -\epsilon_{\mu\nu\rho\sigma}$, puisqu'on abaisse toujours trois indices spatiaux dans cette opération. On a enfin $\epsilon^{\mu\nu\rho\sigma}\epsilon_{\mu\nu\rho\sigma} = -24$ (une somme de 256 termes qui nous fait apprécier à sa juste valeur l'élégance des notations d'Einstein).

2.4 Dérivation et analyse vectorielle

La dernière étape à franchir est de refonder, dans notre formalisme quadridimensionnel, l'analyse vectorielle. Nous commencerons par redéfinir le gradient, pour généraliser ensuite aux autres opérateurs

différentiels. Nous verrons que des lois bien connues de l'électromagnétisme prennent une forme très simple en termes de ces opérateurs. Nous définirons enfin une intégration dans l'espace–temps et généraliserons les théorèmes de Stokes/Ostrogradski.

2.4.1 Dérivation

Nous pouvons définir, pour une fonction du 4–vecteur position d'un événement, la dérivation par rapport aux coordonnées contravariantes de l'événement:

$$\partial_\mu = \frac{\partial}{\partial x^\mu} \ . \tag{2.76}$$

Il est évident que les 4 quantités ∂_μ forment les quatre composantes covariantes d'un opérateur différentiel vectoriel qui généralise la notion de "nabla" à notre espace à quatre dimensions. Leur covariance est manifeste si nous examinons l'accroissement infinitésimal d'un fonction scalaire des x^μ pour un accroissement dx^μ:

$$df = \partial_\mu f(x^\nu) dx^\mu \ . \tag{2.77}$$

df étant un scalaire et dx^μ un vecteur contravariant, ∂_μ est un "vecteur" covariant. Il se transforme donc comme tel dans une transformation de Lorentz:

$$\partial'_\mu = \mathcal{L}_\mu{}^\nu \partial_\nu \ , \tag{2.78}$$

où les ∂' représentent les dérivées par rapport aux nouvelles coordonnées contravariantes. Cette loi de transformation peut aussi s'établir péniblement à partir des lois de transformation des composantes.

On peut aussi définir la dérivation par rapport aux coordonnées covariantes:

$$\partial^\mu = \frac{\partial}{\partial x_\mu}. \tag{2.79}$$

Ces opérateurs différentiels forment évidemment les composantes contravariantes d'un vecteur (comme on peut s'en convaincre en écrivant l'accroissement d'une fonction scalaire des coordonnées covariantes). On a de plus

$$\partial^\mu = g^{\mu\nu} \partial_\nu \ , \tag{2.80}$$

généralisation aux opérateurs différentiels des lois d'élévation ou d'abaissement des indices.

2.4.2 Analyse vectorielle

On peut définir à partir des ∂_μ des analogues des opérateurs de l'analyse vectorielle à trois dimensions. Si f est une fonction scalaire, $\partial_\mu f$, généralisant le gradient, peut s'écrire

$$\partial_\mu f = \left(\frac{1}{c} \frac{\partial f}{\partial t}, \boldsymbol{\nabla} f \right) \tag{2.81}$$

et

$$\partial^\mu f = \left(\frac{1}{c} \frac{\partial f}{\partial t}, -\boldsymbol{\nabla} f \right) \tag{2.82}$$

Si on considère un champ de 4–vecteurs $A^\mu(x^\nu) = (a^0, \mathbf{a})$, on peut définir sa divergence comme:

$$\partial^\mu A_\mu = \partial_\mu A^\mu = \frac{1}{c} \frac{\partial a^0}{\partial t} + \boldsymbol{\nabla} \cdot \mathbf{a} \ , \tag{2.83}$$

qui n'est pas sans nous rappeler l'équation de conservation de la charge ou la jauge de Lorentz. L'analogue du rotationnel sera le tenseur de rang 2 complètement antisymétrique:

$$\partial^\mu A^\nu - \partial^\nu A^\mu \ . \tag{2.84}$$

On pourra aussi écrire le rotationnel sous sa forme complètement covariante:

$$\partial_\mu A_\nu - \partial_\nu A_\mu \ . \tag{2.85}$$

Le "laplacien" de l'espace temps est la norme du vecteur ∂^μ:

$$\partial_\mu \partial^\mu = \frac{1}{c^2}\frac{\partial^2}{\partial t^2} - \triangle = \Box \tag{2.86}$$

et n'est pas autre chose que le d'alembertien $\Box$.

Nous pouvons ainsi réécrire de façon extrêmement simple quelques lois de l'électromagnétisme classique. L'équation de conservation de la charge s'écrit simplement $\partial_\mu J^\mu = 0$, en utilisant le 4–vecteur courant.

Les équations de Poisson pour les potentiels peuvent se regrouper en $\partial_\mu \partial^\mu \mathcal{A}^\nu = \mu_0 J^\nu$, à condition de regrouper les potentiels scalaire et vecteur en un 4–vecteur $\mathcal{A}^\mu = (V/c, \mathbf{A})$. L'équation de Poisson prouve immédiatement que cette quantité est un 4–vecteur. On pourrait alors en déduire les lois relativistes de transformation des potentiels dans un changement de référentiel et, par dérivation, les lois de transformation des champs. Nous établirons plus simplement ces résultats dans les prochains chapitres.

La condition de jauge de Lorentz s'écrit tout simplement $\partial_\mu \mathcal{A}^\mu = 0$. L'ensemble de ces relations prouve que l'électromagnétisme se coule de manière très naturelle dans le cadre mathématique de la cinématique relativiste.

2.4.3 Intégration

On peut définir une intégrale de volume dans l'espace–temps pour n'importe quel type de quantité par

$$\int \quad d\Omega \ , \tag{2.87}$$

où $d\Omega = cdtdxdydz$ est l'élément d'intégration dans l'espace temps dont nous avons déjà démontré le caractère scalaire.

Une surface dans l'espace à trois dimensions est une variété à trois dimensions. On peut définir une intégrale sur ces surfaces (un flux) à condition de définir un 4–vecteur élément de surface dS^μ.

Un élément de surface est un petit objet à trois dimensions. On peut le considérer comme sous-tendu par trois 4–vecteurs dx^μ, dy^μ et dz^μ (comme un petit élément de surface à deux dimensions est sous–tendu par deux vecteurs infinitésimaux). dS^μ doit être orthogonal à tout vecteur de l'élément et sa longueur doit être une mesure du "volume" de l'élément de surface. On peut obtenir dS^μ par une procédure d'orthogonalisation standard. On forme d'abord le tenseur de rang 3 $dS^{\mu\nu\rho}$ tel que:

$$dS^{\mu\nu\rho} = \begin{vmatrix} dx^\mu & dy^\mu & dz^\mu \\ dx^\nu & dy^\nu & dz^\nu \\ dx^\rho & dy^\rho & dz^\rho \end{vmatrix} \tag{2.88}$$

L'élément de surface cherché s'obtient alors en contractant ce tenseur avec le tenseur de rang 4 complètement antisymétrique:

$$dS^\sigma = -\frac{1}{6}\epsilon^{\sigma\mu\nu\rho} dS_{\mu\nu\rho} \ , \tag{2.89}$$

dont on montre qu'il a toutes les propriétés requises. Le lecteur sceptique aura avantage à écrire explicitement le vecteur élément de surface pour un petit élément de volume purement spatial (qui est bien un élément de surface dans l'espace–temps). Il trouvera un 4–vecteur dont seule la composante temporelle est non nulle. Sa mesure est précisément le volume spatial de l'élément considéré.

On peut établir pour les intégrales de surface un théorème généralisant le théorème de gauss:

$$\int_S A^\mu dS_\mu = \int_V \partial_\mu A^\mu \, d\Omega \ , \tag{2.90}$$

où V est un volume dans l'espace–temps et S sa surface frontière. L'intégrale de la divergence étendue à tout l'espace est donc égale au flux sur la "sphère de l'infini"[3]. Celui-ci est en général nul pour des champs physiques.

On peut aussi définir une intégration sur des variétés à deux dimensions (que nous n'appellerons pas surfaces). L'élément d'intégration est un tenseur antisymétrique de rang 2 formé sur les vecteurs dx^μ et dy^ν sous–tendant l'élément d'intégration:

$$df^{\mu\nu} = dx^\mu dy^\nu - dx^\nu dy^\mu \ . \tag{2.91}$$

On peut enfin définir une intégrale curviligne sur une ligne d'univers. Le théorème de Stokes relie l'intégrale sur une variété à deux dimensions à l'intégrale sur son contour:

$$\int A_\mu dx^\mu = \int df^{\mu\nu}(\partial_\mu A_\nu - \partial_\nu A_\mu) \ . \tag{2.92}$$

[3]Cette sphère de l'infini est assez peu intuitive. Elle est formée de tous les points de l'espace à $t = -\infty$ et de tout l'espace à nouveau à $t = \infty$. Entre les deux, elle est formée à tout instant de tous les points à l'infini dans l'espace (la "sphère de l'infini" habituelle de l'analyse vectorielle).

Chapitre 3

Dynamique relativiste

Après avoir jeté les bases d'une nouvelle cinématique, il nous faut bien sûr établir la nouvelle dynamique. Nous commencerons par postuler une forme très simple pour l'action d'une particule libre, dont nous déduirons les équations de Lagrange (qui donnent trivialement un mouvement rectiligne uniforme). Nous en déduirons, de manière plus intéressante, l'expression relativiste de l'impulsion de la particule. Nous définirons alors la notion de force pour une particule en interaction. Nous n'irons pas beaucoup plus loin dans ce chapitre. Pour utiliser la notion de force, il faut en effet la relier aux causes du mouvement. C'est ce que nous ferons au prochain chapitre dans le cas de l'électromagnétisme, en "redécouvrant" l'expression de la force de Lorentz. Notons à ce point que toutes les forces phénoménologiques utilisées en mécanique classique (frottements, tensions..) n'ont aucun sens en relativité (elles ne peuvent se transformer correctement dans un changement de référentiel). Notons également que la force de gravitation ne peut être traitée correctement que dans le cadre de la relativité générale. Nous conclurons ce chapitre par quelques brèves remarques sur le vaste problème des collisions de particules relativistes.

3.1 Particule Libre

Nous essayons ici de définir une action ou un lagrangien pour une particule relativiste libre. La condition d'extrémalité de l'action devrait alors nous donner la ligne d'univers suivie par cette particule.

Pour que toutes les quantités que nous manipulons se comportent bien dans un changement de référentiel (on dit souvent qu'elles sont manifestement covariantes), il faut que l'action et le lagrangien soient tous les deux des 4–scalaires. Le temps sur lequel on intègre le lagrangien doit donc être lui aussi un 4–scalaire. De manière évidente, seul le temps propre τ de la particule convient. Enfin, les bornes de l'intégration doivent être deux événements bornant la portion de ligne d'univers cherchée (et jouant le rôle de conditions aux limites dans les équations de Lagrange). On écrira donc:

$$S = \int_a^b L\, d\tau \ , \tag{3.1}$$

ou encore:

$$S = \frac{1}{c} \int_a^b L\, ds \ , \tag{3.2}$$

où ds est un intervalle infinitésimal sur la ligne d'univers: $ds = cd\tau$.

Pour L, le choix le plus simple est une constante $-\alpha$, relative à la particule. Nous choisissons un signe moins et une constante positive. Ce choix conduira à un minimum pour l'action[1]. On peut

[1] On peut rendre l'intégrale du temps propre aussi petite que l'on veut en imaginant des mouvements à de très grandes vitesses entre a et b. En revanche, l'intégrale du temps propre est maximale si la particule se déplace à vitesse constante entre a et b. L'action, pour avoir un minimum, doit donc être opposée à l'intégrale du temps propre.

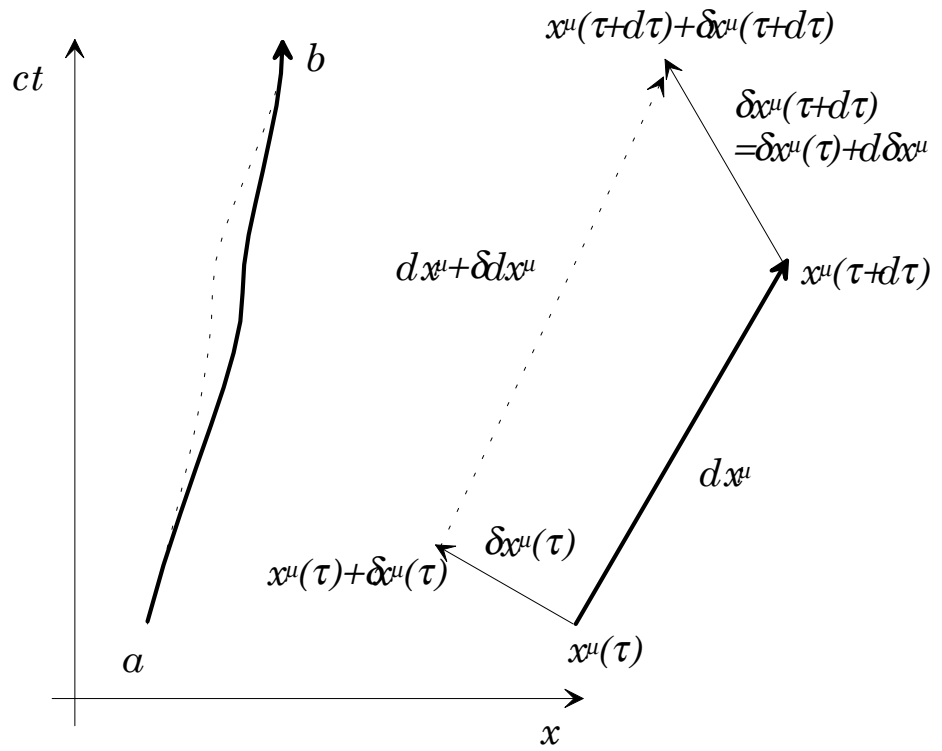

Figure 3.1: Deux lignes d'univers joignant deux événements. L'une est la trajectoire de référence (en traits pleins), l'autre une trajectoire infiniment proche (en pointillés). Les deux trajectoires coïncident pour les événements initial et final. La partie de droite schématise des éléments infinitésimaux correspondants des deux trajectoires.

identifier aisément cette constante en imposant à notre action de coïncider avec l'action classique si la vitesse de la particule est petite devant c. En remarquant que $d\tau = dt/\gamma$, on écrira:

$$S = -\alpha \int \sqrt{1 - v^2/c^2}\, dt = \int (-\alpha + \alpha v^2/2c^2)\, dt \ . \tag{3.3}$$

Le terme constant ne joue aucun rôle dans l'action classique. Pour que le terme en v^2 coïncide avec l'énergie cinétique, il faut que $\alpha = mc^2$. Le lagrangien d'une particule libre est donc une simple constante:

$$L = -mc^2 \tag{3.4}$$

et l'action s'écrit simplement:

$$S = -mc^2 \int d\tau = -mc \int ds \ . \tag{3.5}$$

Nous allons maintenant établir, à partir de cette expression de l'action, les équations de Lagrange. Ce raisonnement est très proche de celui utilisé dans la première partie de ce cours pour établir les équations de Lagrange d'un système quelconque. Nous considérerons donc deux lignes d'univers légèrement différentes entre les événements a et b qui nous serviront de "conditions aux limites" (voir figure 3.1). Nous ne pouvons plus en effet spécifier, sans perdre l'invariance relativiste, de positions à un instant initial et à un instant final. L'une de ces trajectoires sera la trajectoire effectivement suivie, qui réalise un extremum pour l'action. Elle est paramétrée par le temps propre τ de la particule. Les événements de cette ligne d'univers peuvent donc s'écrire $x^\mu(\tau)$.

L'autre ligne d'univers s'écarte de la trajectoire de référence par des quantités infinitésimales. Les coordonnées spatio–temporelles de chaque événement sur cette ligne d'univers pourront s'écrire $x^\mu(\tau) + \delta x^\mu(\tau)$, où $\delta x^\mu(\tau)$ est une quantité infinitésimale. Notons qu'avec ce choix d'écriture, la

trajectoire variée et la trajectoire de référence sont paramétrées par le même temps propre τ. C'est en effet une condition importante pour pouvoir écrire les intégrales d'action avec le même élément différentiel. Notons que le temps propre est utilisé comme un simple paramètre. τ n'est pas le temps propre sur la trajectoire variée (c'est précisément à la détermination de celui-ci que nous allons nous consacrer).

Considérons un intervalle infinitésimal de temps propre $d\tau$ (voir partie droite de la figure 3.1). Sur la trajectoire de référence, on passe de l'événement x^μ à $x^\mu + dx^\mu$ avec $dx^\mu = (dx^\mu/d\tau)d\tau$. Sur la ligne d'univers variée, en revanche, on passe de l'événement $x^\mu + \delta x^\mu$ à $x^\mu + \delta x^\mu + dx^\mu + d\delta x^\mu$ avec $d\delta x^\mu = (d\delta x^\mu/d\tau)d\tau = \delta(dx^\mu/d\tau)d\tau = \delta dx^\mu$, l'accroissement de la vitesse étant manifestement la dérivée temporelle de l'accroissement. Dans toute la suite du raisonnement, nous traiterons les accroissements δ comme des infiniments petits d'ordre supérieur par rapport aux éléments d de trajectoire. L'action sur la trajectoire de référence est donc

$$S = -mc \int_a^b ds = -mc \int_a^b \sqrt{dx_\mu dx^\mu} \ . \tag{3.6}$$

L'action sur la trajectoire variée s'écrit:

$$S + \delta S = -mc \int_a^b \sqrt{(dx_\mu + \delta dx_\mu)(dx^\mu + \delta dx^\mu)} \ . \tag{3.7}$$

En développant cette dernière expression au premier ordre non nul dans les petits accroissements, on obtient l'accroissement de l'action

$$\begin{aligned} S + \delta S &= -mc \int_a^b \left[1 + \frac{dx_\mu \delta dx^\mu}{ds^2}\right] ds \\ \delta S &= -mc \int_a^b \frac{dx_\mu}{ds} \delta dx^\mu \\ &= -m \int_a^b U_\mu \delta dx^\mu \\ &= \ , -m \int_a^b U_\mu \frac{d\delta x^\mu}{d\tau} d\tau \end{aligned} \tag{3.8}$$

où nous faisons intervenir la 4–vitesse $U_\mu = c dx_\mu/ds$ de la particule.

Pour obtenir les équations de Lagrange, nous pouvons transformer cette expression par une intégration par parties:

$$\delta S = -m \int_a^b U_\mu \delta dx^\mu = -m \left[U_\mu \delta x^\mu\right]_a^b + m \int_a^b \frac{dU_\mu}{d\tau} \delta x^\mu d\tau \ . \tag{3.9}$$

Le terme tout intégré est manifestement nul puisque les deux lignes d'univers coïncident sur les événements limites. La nullité de δS ne peut être assurée pour tous les accroissements que si:

$$\frac{dU^\mu}{d\tau} = 0 \ , \tag{3.10}$$

ce qui constitue l'équation de Lagrange, nécessairement triviale, du mouvement de la particule libre.

En fait, le principal intérêt de ce calcul ne réside pas dans cette équation que nous aurions pu obtenir en invoquant l'invariance galiléenne. Nous avons d'abord pris la mesure des difficultés mises en jeu dans un calcul variationnel en relativité (en particulier, nous avons du prendre garde de ne manipuler que des quantités 4–vectorielles ou scalaires). Ensuite, nous avons établi, en passant, l'équation (3.9) qui va nous permettre de déterminer les variations de l'action dans une variation du point d'arrivée, et donc d'établir la forme relativiste de l'impulsion.

3.2 Energie–impulsion

Nous avions introduit naturellement l'énergie et l'impulsion en dynamique classique comme les dérivées temporelles et spatiales de l'action par rapport au point d'arrivée. La contrepartie relativiste en est manifestement le 4–gradient de l'action par rapport à l'événement b. Nous poserons donc:

$$P_\mu = -\partial_\mu S \ , \tag{3.11}$$

les dérivées s'entendant par rapport à la position de b.

Pour estimer P_μ, nous considérerons deux lignes d'univers effectivement suivies par la particule, l'une (ligne d'univers de référence) connectant a et b, l'autre (variée) a et $b+db$. Le calcul de la variation de l'action suit alors celui effectué au paragraphe précédent. Les deux trajectoires sont paramétrées par le temps propre de la trajectoire de référence. La seule différence est que l'accroissement ne s'annule pas pour l'événement b: $\delta x^\mu(b) = db$. L'équation (3.9) est donc encore correcte. Comme les deux trajectoires sont des trajectoires effectivement suivies, elles correspondent à des vitesses constantes et le terme intégral est identiquement nul. En revanche, le terme tout intégré est non nul et vaut $U_\mu \delta x^\mu(b)$. On en déduit par simple comparaison que:

$$P_\mu = mU_\mu \qquad P^\mu = mU^\mu \ , \tag{3.12}$$

un résultat pour le moins attendu.

Penchons nous maintenant sur la signification physique des composantes de P^μ. On a bien sûr $P^\mu = (-(1/c)dS/dt, \boldsymbol{\nabla} S)$. Une simple comparaison avec la mécanique classique nous indique que la composante temporelle de P^μ est l'énergie mécanique $\mathcal{E}$ de la particule, alors que les composantes spatiales représentent la quantité de mouvement:

$$P^\mu = (\mathcal{E}/c, \mathbf{p}) \ . \tag{3.13}$$

L'énergie et la quantité de mouvement apparaissent donc comme les composantes temporelles et spatiales d'un 4–vecteur. Elles doivent donc se transformer ensemble dans un changement de référentiel. Nous pouvons de plus identifier ces quantités à partir de l'expression de la 4–vitesse:

$$\mathcal{E} = m\gamma c^2 \tag{3.14}$$

$$\mathbf{p} = m\gamma \mathbf{v} \tag{3.15}$$

Commentons tout d'abord l'expression de l'énergie mécanique. Pour des mouvements à vitesse faible devant c, on peut développer le facteur γ. On obtient alors:

$$\mathcal{E} = mc^2 + \frac{1}{2}mv^2 \ . \tag{3.16}$$

A une constante près, nous retrouvons donc bien l'énergie cinétique galiléenne.

Si les énergies sont toujours définies à une constante additive près en mécanique classique, il n'en est pas de même en relativité. On ne saurait en effet ajouter une constante arbitraire à la composante temporelle d'un 4–vecteur en conservant une forme correcte pour les lois de changement de référentiel. La constante additive qui, ci-dessus, représente l'énergie mécanique d'une particule au repos:

$$\mathcal{E}_0 = mc^2 \ , \tag{3.17}$$

doit donc avoir une signification physique[2].

Si une masse au repos possède de l'énergie, cela indique qu'on peut, au moins sur un plan de pur bilan d'énergie, transformer une certaine quantité de masse en énergie ou une certaine quantité

[2]Remarquons que nous venons d'établir ce qui est sans doute la formule de physique la plus célèbre.

d'énergie en masse. Les applications les plus spectaculaires de ces conversions sont les créations et annihilations de particules élémentaires. Quand deux antiparticules s'annihilent en émettant de l'énergie sous forme de rayonnement électromagnétique, on a une conversion totale d'une certaine quantité de masse en énergie. Le phénomène réciproque, la création d'une paire particule/antiparticule à partir d'un rayonnement suffisamment énergétique, correspond à la conversion d'énergie en masse. Pour créer, par exemple, une paire électron/positron, il faudra au moins fournir une énergie $2mc^2$, soit 2 fois 511 keV. Cette notion de seuil de réaction joue un rôle essentiel pour l'analyse des collisions de particules. Notons à ce point que, si nous avons montré que de tels processus sont envisageables du point de vue du simple bilan énergétique, nous n'avons pas montré qu'ils pouvaient se produire. Il faut ajouter à ces conditions de bilan les règles de sélection qui indiquent quelles particules on peut effectivement créer dans une situation donnée. Ces règles ne peuvent s'obtenir que par une approche quantique qui est hors de propos ici. En dépit des conséquences très importantes de cette énergie de repos pour la physique des particules, nous n'évoquerons pas ce problème beaucoup plus avant ici, et nous renverrons le lecteur intéressé aux nombreux manuels qui couvrent ce sujet.

La quantité de mouvement, de son côté, s'écrit donc $m\gamma\mathbf{v}$. Au premier ordre en v/c, elle coïncide donc bien avec la quantité de mouvement galiléenne ordinaire. En revanche, pour des vitesses proches de celles de la lumière, la quantité de mouvement croît beaucoup plus vite que la vitesse (elle est en fait simplement proportionnelle à la célérité). Comme l'accroissement de la quantité de mouvement est manifestement relié à la force qui s'exerce sur la particule (nous anticipons quelque peu sur la suite de cet exposé), on retrouve qu'il est impossible d'accélérer une particule matérielle jusqu'à la vitesse de la lumière, puisqu'il faudrait lui communiquer une quantité de mouvement infinie. Notons enfin que, du point de vue de la quantité de mouvement, et dans une large mesure du point de vue de l'ensemble de la dynamique, "tout se passe comme si" la masse de la particule dépendait de la vitesse comme le facteur γ. Cette approche, largement développée dans des ouvrages de vulgarisation, n'est bien sûr pas correcte. L'invariance relativiste impose que la masse de la particule soit un 4–scalaire, indépendant du référentiel.

Notons, pour terminer ce paragraphe, un lien utile entre énergie et quantité de mouvement. Nous savons en effet que le module de la 4–vitesse est égal à c. On en déduit $P_\mu P^\mu = m^2c^2$ et donc

$$p^2c^2 + m^2c^4 = \mathcal{E}^2 \ . \tag{3.18}$$

Examinons le cas d'une particule ultra–relativiste, dont la vitesse est proche de celle de la lumière. L'énergie $\mathcal{E}$ est alors très supérieure à l'énergie de repos $\mathcal{E}_0$, et on peut négliger le terme m^2c^4 dans l'expression précédente. On a alors simplement $\mathcal{E} = pc$ et on peut pratiquement, à une constante dimensionnelle près, confondre l'énergie et l'impulsion de la particule.

Un exemple extrême de particule ultra–relativiste est une particule de masse nulle, comme le photon ou le neutrino (si la masse du premier est très certainement nulle, un petit doute théorique et expérimental subsiste pour le second). Si la vitesse de propagation de cette particule était inférieure à c, le facteur γ étant fini, l'énergie de la particule serait nulle. Pour que de telles particules aient une énergie non nulle, il faut que leur vitesse de propagation soit la vitesse limite de la relativité (qu'on peut donc confondre à ce point seulement, en toute rigueur, avec la vitesse de la lumière). Par un simple passage à la limite, on voit donc que l'énergie et l'impulsion de la particule sont liées par $\mathcal{E} = pc$. Dans le cas particulièrement important du photon, l'énergie est $\mathcal{E} = h\nu$, où h est la constante de Planck et l'impulsion est donc, en module, $h\nu/c$. Si on y ajoute que l'impulsion doit avoir la direction du vecteur d'onde (nous verrons pourquoi à la fin de cette partie), on en déduit que l'impulsion spatiale du photon peut s'écrire $\mathbf{p} = \hbar\mathbf{k}$.

3.3 Particule soumise à une force

Cette section ne peut être à ce niveau qu'embryonnaire. Si la particule est soumise à des interactions qui modifient son mouvement, on pourra écrire l'équation du mouvement sous la forme d'une

équation de Lagrange à condition de connaître le lagrangien décrivant l'interaction. Le seul cas où nous pourrons effectuer cette démarche est celui de l'électromagnétisme que nous traiterons dans le prochain chapitre. En fait, aucune autre interaction ne peut être incluse correctement dans le cadre de la relativité restreinte à notre niveau. Tout d'abord, toutes les interactions phénoménologiques (contacts, ressorts..) n'ont plus de sens en relativité, ainsi que la notion de solide. Elles contiennent en effet toutes plus ou moins la notion d'interaction instantanée. Il est possible de formuler une hydrodynamique relativiste, par exemple, mais au prix de grands efforts. Même la gravitation ne peut être en toute rigueur coulée dans ce cadre. Elle n'est décrite convenablement que dans celui de la relativité générale.

En fait, nous ne pouvons ici qu'intuiter que la dérivée par rapport au temps propre de l'impulsion, $G^\mu = dP^\mu/d\tau$, jouera le rôle d'une force, qu'on pourra exprimer simplement, dans le cas de l'électromagnétisme, en fonction de la vitesse de la particule et des champs (cette force ne sera autre que la force de Lorentz). Le module de l'impulsion étant constant, la force est nécessairement perpendiculaire à l'impulsion: $G_\mu P^\mu = 0$. En termes de l'énergie et de la quantité de mouvement spatiale, on pourra écrire:

$$G^\mu = (\frac{\gamma}{c}\frac{d\mathcal{E}}{dt}, \gamma\frac{d\mathbf{p}}{dt}) \ . \tag{3.19}$$

On pourra poser $\mathbf{f} = d\mathbf{p}/dt$ et définir ainsi la force spatiale, dérivée par rapport au temps ordinaire de la quantité de mouvement. L'orthogonalité de G_μ et de P^μ est simplement équivalente, comme on s'en convaincra aisément, au fait que $\mathcal{E}^2 - c^2p^2$ est une constante.

3.4 Conservation de l'énergie–impulsion. Application aux collisions

On peut montrer, comme en mécanique classique, que l'invariance dans une translation globale dans l'espace–temps implique que tout système isolé possède un 4–vecteur énergie impulsion qui est conservé. Pour un ensemble de particules matérielles sans interaction, l'impulsion globale du système est simplement la somme des impulsions individuelles. Pour des particules en interaction (en particulier électromagnétique), la situation est moins simple. L'interaction, qui ne peut être instantanée, doit en effet être véhiculée par un champ. Ce champ, possédant une énergie, doit aussi posséder une quantité de mouvement et entrer dans le bilan de l'impulsion globale. L'analyse détaillée de la conservation de l'impulsion dans une telle situation est alors difficile.

La conservation de l'impulsion permet néanmoins d'obtenir des renseignements précis sur les collisions de particules relativistes. Dans une telle collision, on considère en effet un état initial où les deux particules sont très éloignées et n'interagissent pratiquement pas. Après avoir interagi, les deux particules s'éloignent à nouveau l'une de l'autre et on considère un état final où, à nouveau, l'interaction est négligeable. Même si on renonce à examiner ce qui se passe pendant l'interaction, on pourra écrire que l'impulsion globale initiale, qui est la somme des impulsions de particules incidentes, est égale à la somme des impulsions finales. On regroupe ainsi dans une même équation liant deux 4–vecteurs les lois de conservation de l'énergie et de l'impulsion qu'on écrit séparément dans une analyse classique de la collision[3].

Pour illustrer très brièvement l'étude des collisions relativistes, nous traiterons deux exemples: le calcul d'un seuil de réaction et le cas de l'effet Compton. Ces deux exemples à eux seuls sont loin de

[3]Si les particules sont chargées, un champ électromagnétique est associé à chaque particule avant et après la collision. Ce champ doit entrer en toute rigueur dans le bilan d'impulsion. En fait, quand on mesure la masse ou l'impulsion d'une particule chargée, on ne le fait jamais indépendamment du champ qui accompagne cette particule. La masse que nous mettons dans l'impulsion est donc une masse "habillée", qui englobe effectivement la contribution du champ à l'inertie de la particule. Nous n'aurons donc effectivement pas à tenir compte des champs des particules libres dans notre bilan d'impulsion. Notons qu'avec une particule chargée ponctuelle, la différence entre la masse "nue" et la masse "habillée" de la particule est évidemment infinie. Ce n'est qu'au prix d'un calcul complètement quantique mené avec les techniques de renormalisation qu'on peut définir proprement la masse "habillée" de la particule.

couvrir toutes les applications de la conservation de la 4–impulsion à la physique des particules, mais ils permettent d'illustrer de façon simple des techniques et des concepts importants.

3.4.1 Seuil de réaction

La plupart des collisions réalisées dans les accélérateurs visent à produire de nouvelles particules. Il est clair que l'énergie cinétique incidente doit être suffisante pour que la réaction soit énergétiquement possible. Le seuil de réaction est précisément l'énergie cinétique à fournir aux particules en collision pour que la réaction soit énergétiquement possible. Ceci ne signifie pas nécessairement que la réaction se produise effectivement. Des règles de sélection quantiques ou une section efficace insuffisante peuvent faire que les produits ne soient jamais observés.

Le calcul du seuil est complexe pour les réactions complexes. Nous ne considérerons donc que l'une des plus simples des réactions de création de particules: la création d'une paire proton/antiproton par collision de deux protons. L'équation de la réaction s'écrit:

$$p + p \longrightarrow p + p + p + \overline{p} \ . \tag{3.20}$$

L'énergie minimale pour que cette réaction puisse se produire correspond à une situation où toutes les particules finales seraient au repos. Elle est donc définie dans un référentiel où la quantité de mouvement totale est nulle. Pour tout système de particules matérielles, il existe un tel référentiel, appelé "référentiel du centre de masse" $\mathcal{R}_{CM}$. Il est l'analogue du référentiel barycentrique en mécanique classique. Dans $\mathcal{R}_{CM}$, l'énergie minimale est donc $4mc^2$ (les particules et les antiparticules ont même masse m). Dans ce référentiel, les deux protons incidents ont initialement des énergies égales et des quantités de mouvement opposées. L'énergie de chacun doit donc être $2mc^2$ et son énergie cinétique mc^2. Le seuil de réaction, dans le référentiel barycentrique, est donc de mc^2 pour chaque particule, un résultat assez intuitif.

Le problème est que l'énergie à fournir effectivement est celle définie dans le référentiel du laboratoire, qui ne coïncide pas nécessairement avec le référentiel du centre de masse. Ce n'est que dans le cas des anneaux de collision (LEP pour les collisions électron/électron, par exemple), que les deux projectiles sont de quantités de mouvement opposées et que le référentiel du laboratoire est aussi celui du centre de masse.

Dans beaucoup d'expériences, plus anciennes, un seul des protons est en mouvement (le projectile) et l'autre (la cible) est immobile. On peut estimer le seuil de réaction dans ce cas. La première chose à faire est de déterminer les éléments de $\mathcal{R}_{CM}$ par rapport au référentiel du laboratoire $\mathcal{R}$. Nous noterons $\mathcal{E}$ et p l'énergie et la quantité de mouvement du projectile. Tous les mouvements s'effectuant sur l'axe projectile/cible, nous n'écrirons que des quantités en projection sur cet axe. Il est clair aussi que la vitesse de $\mathcal{R}_{CM}$ par rapport à $\mathcal{R}$ est selon cet axe. Le 4–vecteur énergie–impulsion total avant la collision s'écrit donc $(\mathcal{E}/c + mc, p)$. De manière évidente la quantité de mouvement totale n'est pas nulle.

Nous cherchons donc une transformation de Lorentz (paramètres β et γ) telle que la nouvelle impulsion p' soit nulle. Il faut donc choisir:

$$\beta = \frac{pc}{\mathcal{E} + mc^2} \ , \tag{3.21}$$

ce qui définit le mouvement du référentiel du centre de masse. On peut alors écrire l'énergie totale disponible dans $\mathcal{R}_{CM}$ en utilisant la même transformation de Lorentz:

$$\mathcal{E}' = \gamma(\mathcal{E} + mc^2 - \beta pc) \ . \tag{3.22}$$

En y portant l'expression de β et en notant que $\mathcal{E}^2 - p^2c^2 = m^2c^4$, on a simplement:

$$\mathcal{E}' = 2\gamma mc^2 \ . \tag{3.23}$$

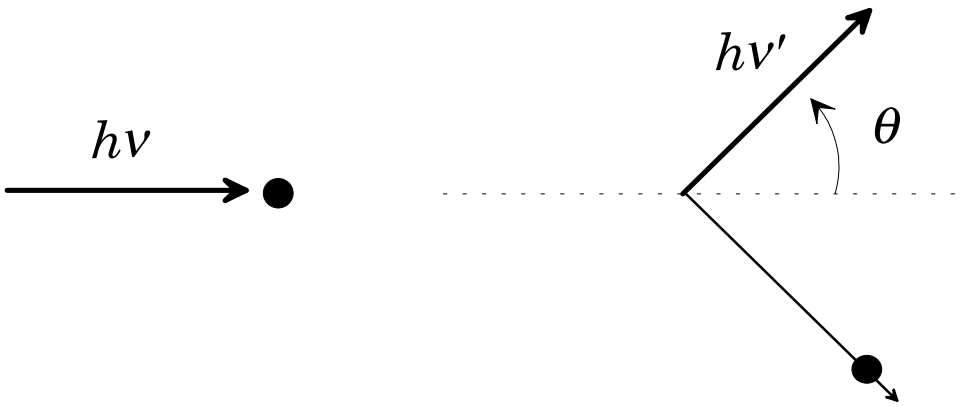

Figure 3.2: Diffusion d'un photon par un électron initialement au repos. Partie gauche: état initial. Partie droite: état final. Le photon est diffusé avec une fréquence modifiée et l'électron est éjecté.

On peut enfin exprimer γ en fonction de β puis de $\mathcal{E}$ et p:

$$\gamma = \frac{1}{\sqrt{1-\beta^2}} = \frac{\sqrt{\mathcal{E}+mc^2}}{\sqrt{2}\sqrt{mc^2}} \ . \tag{3.24}$$

L'énergie disponible dans $\mathcal{R}_{CM}$ est donc simplement:

$$\mathcal{E}' = \sqrt{2mc^2(\mathcal{E}+mc^2)} \ . \tag{3.25}$$

Le seuil de réaction s'obtient alors en écrivant que $\mathcal{E}'$ doit être plus grand que $4mc^2$. Après quelques manipulations, on voit que le seuil s'écrit:

$$\mathcal{E} > 7mc^2 \ , \tag{3.26}$$

l'énergie cinétique du projectile devant être d'au moins $6mc^2$. On doit comparer ce résultat à celui obtenu quand le référentiel du laboratoire est aussi celui du centre de masse. Au lieu d'une énergie cinétique par particule de mc^2, on a besoin d'une énergie 6 fois plus grande. Avec une puissance donnée d'accélérateur, il est donc, de beaucoup, préférable de travailler avec des collisions entre deux particules en mouvement. Ceci explique que pratiquement toutes les grandes machines sont maintenant des anneaux de collision. Le prix à payer est bien sûr que la densité du faisceau est très petite devant celle d'une cible solide: le taux de collisions est beaucoup moins grand que ce qu'on obtient en envoyant un faisceau unique sur une cible fixe.

3.4.2 Effet Compton

Notre second exemple sera l'effet Compton, la diffusion d'un photon de haute énergie par un électron initialement immobile (ou lié à un atome: les énergie mises en jeu dans les expériences sont telles, comme nous le verrons, que l'énergie de liaison d'un électron dans un atome est négligeable par rapport à l'énergie finale de l'électron diffusé). Les paramètre importants du problème sont représentés sur la figure 3.2.

Sans restreindre la généralité, le photon de fréquence ν et d'énergie $h\nu$ est incident dans la direction de l'axe Ox. Après la collision, l'électron et le photon sont diffusés dans des directions qui dépendent des détails de l'interaction. Toujours sans restreindre la généralité, nous pouvons supposer que les trajectoires finales de l'électron et du photon sont dans le plan Oxy. L'électron emportant de l'énergie cinétique, le photon perd nécessairement de l'énergie. Sa nouvelle fréquence sera notée ν'. Nous chercherons seulement ici à calculer la nouvelle fréquence du photon en fonction de son angle de

diffusion θ. C'est en effet la seule quantité accessible dans les expériences: l'électron est rapidement diffusé et son énergie amortie par le milieu où se produit la diffusion. De façon très remarquable, la simple application de la conservation de l'impulsion relativiste permet d'expliciter ce calcul[4].

Les 4–impulsions du photon avant et après la collision seront notées $\mathbf{P}$ et $\mathbf{P}'$. On notera de même $\mathbf{Q}$ et $\mathbf{Q}'$ les impulsions initiale et finale de l'électron. En nous souvenant que la quantité de mouvement spatiale de l'électron est $h\nu/c$, nous pourrons écrire:

$$\begin{aligned}
\mathbf{P} &= (h\nu/c, h\nu/c, 0) \ , && (3.27)\\
\mathbf{P}' &= (h\nu'/c, h\nu'/c \ \cos\theta, h\nu'/c \ \sin\theta) && (3.28)\\
\mathbf{Q} &= (mc, 0, 0) && (3.29)\\
\mathbf{Q}' &= (\mathcal{E}/c, q_x, q_y) \ . && (3.30)
\end{aligned}$$

Dans les équations précédentes, nous n'avons écrit que les composantes spatiales dans le plan Oxy. $\mathcal{E}$ et q_x, q_y représentent l'énergie et la quantité de mouvement de l'électron diffusé. Pour calculer ν' en fonction de θ, nous allons chercher à éliminer ces inconnues. La loi de conservation de l'énergie–impulsion s'écrit:

$$\mathbf{P} + \mathbf{Q} = \mathbf{P}' + \mathbf{Q}' \ . \tag{3.31}$$

On en déduit immédiatement:

$$\begin{aligned}
(\mathbf{P} - \mathbf{P}')^2 &= (\mathbf{Q} - \mathbf{Q}')^2 && (3.32)\\
\mathbf{P}^2 + \mathbf{P}'^2 - 2\mathbf{P}\cdot\mathbf{P}' &= \mathbf{Q}^2 + \mathbf{Q}'^2 - 2\mathbf{Q}\cdot\mathbf{Q}' \ . && (3.33)
\end{aligned}$$

Le photon étant une particule de masse nulle, $\mathbf{P}^2 = \mathbf{P}'^2 = 0$. De plus $\mathbf{Q}^2 = \mathbf{Q}'^2 = m^2c^2$ où m est la masse de l'électron. Enfin, dans le produit scalaire $\mathbf{Q}\cdot\mathbf{Q}'$, le produit des composantes spatiales est nul et il ne reste que le produit des composantes temporelles: $\mathbf{Q}\cdot\mathbf{Q}' = -2m\mathcal{E}$. On en déduit:

$$\frac{h^2\nu\nu'}{c^2}(1 - \cos\theta) = m(\mathcal{E} - mc^2) \ . \tag{3.34}$$

La simple conservation de l'énergie assure alors que:

$$h\nu + mc^2 = h\nu' + \mathcal{E} \ , \tag{3.35}$$

d'où on déduit finalement

$$\lambda' - \lambda = \lambda_c(1 - \cos\theta) \ , \tag{3.36}$$

$\lambda = c/\nu$ et $\lambda' = c/\nu'$ étant les longueurs d'ondes associées au photon avant et après la collision. La longueur λ_c, appelée longueur d'onde de Compton de l'électron, vaut:

$$\lambda_c = \frac{h}{mc} = 2.10^{-11} \text{ m} \ . \tag{3.37}$$

La très petite valeur numérique de cette quantité fait que l'effet n'est notable que pour des photons incidents de courte longueur d'onde, c'est à dire de très haute énergie. Pour des photons visibles, par exemple, la modification de fréquence due au "recul" de l'électron est tout à fait négligeable. Quand nous étudierons la diffusion de rayonnement par un atome nous négligerons complètement cet effet.

[4]Notons que la mise en évidence de l'effet Compton et l'accord quantitatif avec la loi que nous allons démontrer a constitué, sans doute, une des premières preuves indiscutables de l'existence du photon. L'effet photoélectrique, s'il est parfaitement explicable en termes de photons, peut en effet être décrit dans une théorie semi–classique couplant un champ électromagnétique classique à un détecteur quantifié.

Chapitre 4

Electromagnétisme relativiste

Nous désirons traiter maintenant de façon relativiste l'interaction entre particules chargées par l'intermédiaire d'un champ électromagnétique. Nous pourrions bien sûr supposer connues les équations de Maxwell et montrer directement qu'elles s'insèrent sans difficultés dans un cadre relativiste. Nous allons en fait procéder d'une façon beaucoup moins directe, mais plus profonde. Nous allons chercher à écrire l'interaction relativiste non triviale entre des particules par l'intermédiaire d'un champ de vecteurs. En d'autres termes, nous allons tenter de construire, dans le cadre de la relativité, la théorie de champ la plus simple qui ne soit pas triviale. Nous postulerons pour cela la forme des actions associées aux particules et au champ. Nous verrons que ces formes sont en effet tout à fait naturelles. Nous déduirons alors de ces actions des équations de Lagrange qui décrivent la dynamique des particules couplées au champ et la dynamique du champ couplé aux particules. Nous constaterons enfin que ces équations ont la forme des équations de Maxwell. Nous aurons donc pu déduire les équations de Maxwell d'une approche lagrangienne très générale.

Nous aurons montré également que l'électromagnétisme est la plus simple des théories de champ vectorielles dans le cadre de la relativité. Bien sûr, nous ne nous contenterons pas dans ce chapitre de réécrire des équations bien connues. L'écriture en termes de quantités explicitement covariantes nous fournira des résultats nouveaux, difficiles ou impossibles à établir dans le cadre de l'électromagnétisme classique. Nous établirons ainsi la transformation des champs électromagnétiques dans un changement de référentiel, nous isolerons des invariants scalaires formés à partir des champs, dont nous montrerons qu'ils ont une signification physique importante. Nous pourrons enfin établir des bilans d'énergie–impulsion pour le champ qui nous permettront de jeter une lumière nouvelle sur des phénomènes bien connus, comme la pression de radiation ou même la force de Coulomb électrostatique.

Notre système sera donc constitué d'un ensemble de particules en interaction avec un champ représenté par un champ de 4–vecteurs. L'action totale pour ce système peut a priori se décomposer sous la forme:

$$S = S_{\text{Particules Libres}} + S_{\text{Champ Libre}} + S_{\text{Interaction}} \ , \tag{4.1}$$

où $S_{\text{Particules Libres}}$ représente l'action des particules en l'absence de champ (une simple collection de particules libres si on néglige toute autre forme d'interaction), $S_{\text{Champ Libre}}$ représente l'action décrivant le champ seul, en l'absence de toutes particules. Enfin, $S_{\text{Interaction}}$ représente l'interaction entre particules et champ: d'une part les particules sont la source du champ, d'autre part la présence du champ modifie la trajectoire des particules.

Pour aborder ce problème, nous allons procéder en deux temps, comme on le fait souvent dans les exposés élémentaires d'électromagnétisme. Nous allons d'abord considérer une particule unique en présence d'un champ imposé. Nous supposons donc qu'un grand ensemble de particules crée un champ qui agit sur une particule test. Si cette particule "test" est suffisamment petite, elle ne modifie pas notablement le champ ni la dynamique des particules "sources". Les seules variables dynamiques dans ce cas sont donc celles de la particule libre, les valeurs du champ étant des quantités imposées.

Nous pourrons alors écrire aisément les équations de Lagrange, dont nous verrons qu'elles redonnent l'expression attendue pour la force de Lorentz. Nous verrons en effet que le seul champ de vecteurs introduit (qui n'est autre que la version relativiste du potentiel vecteur) intervient dans la force sous la forme de son rotationnel, c'est à dire d'un tenseur de rang deux antisymétrique. Ce tenseur pouvant être exprimé à partir de deux champs de vecteurs, nous retrouverons que l'électromagnétisme est une théorie à deux champs. En écrivant les propriétés du rotationnel relativiste, nous obtiendrons des relations entre ces deux champs qui se trouveront coïncider avec le groupe des équations de Maxwell homogènes. Notons enfin que nous pourrons dès ce point établir la forme des changements de référentiels pour les champs.

Dans un deuxième temps, nous considérerons un champ en interaction avec des particules dont la dynamique est imposée. Nous représenterons les mouvements de ces particules par des densités de charges et de courants et nous postulerons une forme simple pour $S_{\text{Champ Libre}}$. Nous obtiendrons alors, comme équations de Lagrange, les équations de Maxwell faisant intervenir les sources. Nous aurons alors terminé notre programme: en combinant les équations décrivant la dynamique des particules dans un champ imposé et les équations décrivant la dynamique du champ sous l'action de courants imposés, on peut, au moins en principe, résoudre tout problème d'électromagnétisme. La dernière partie de ce chapitre sera alors consacrée à l'exploitation de ces résultats. Nous y établirons en particulier les bilans d'énergie–impulsion pour le champ.

4.1 Particule libre dans un champ imposé

4.1.1 Equations de Lagrange

Nous considérons donc ici une particule de masse m plongée dans un champ imposé. L'action décrivant la particule libre s'écrit simplement

$$S_{\text{Particules Libres}} = -mc \int_a^b ds \ , \tag{4.2}$$

où a et b sont deux événements décrivant les conditions aux limites imposées à la particule. En l'absence de champ, la ligne d'univers de la particule serait simplement la droite joignant a et b.

Nous postulerons que le champ peut être représenté par un champ unique de 4–vecteurs que nous noterons $\mathcal{A}^\mu(x^\nu) = (V/c, \mathbf{A})$. Pour des raisons qui apparaîtront évidentes plus tard, nous nommerons le champ $\mathcal{A}$ "potentiel". L'interaction entre le champ et la particule doit être représentée par l'intégrale d'une quantité scalaire sur la ligne d'univers entre les événements limites a et b. La quantité la plus simple non triviale que nous puissions former est donc:

$$S_{\text{Interaction}} = -q \int_a^b \mathcal{A}_\mu \, dx^\mu \ , \tag{4.3}$$

où q est une quantité scalaire représentant l'intensité du couplage de la particule au champ que nous nommerons simplement "charge". Dans l'expression de l'action d'interaction, il faut comprendre, comme au chapitre précédent, que la ligne d'univers de la particule est paramétrée par son temps propre τ et que dx^μ est en fait égal à $(dx^\mu/d\tau)d\tau$.

Nous pouvons tout de suite nous rassurer sur la pertinence de ce lagrangien d'interaction. En écrivant que $(dx^\mu/d\tau)d\tau = U^\mu d\tau = U^\mu dt/\gamma$, en particularisant pour un instant un référentiel $\mathcal{R}$ et en développant le produit scalaire, on met l'action d'interaction sous la forme:

$$S_{\text{Interaction}} = -q \int_a^b (V - \mathbf{v} \cdot \mathbf{A}) \, dt \ , \tag{4.4}$$

où $\mathbf{v}$ est la vitesse tridimensionnelle de la particule dans $\mathcal{R}$. On retrouve bien là la forme du lagrangien d'interaction avec une particule chargée obtenu dans le premier chapitre de la première partie de ce

cours, à condition bien sûr d'assimiler la composante temporelle du potentiel au potentiel scalaire de l'électromagnétisme et ses composantes spatiales au potentiel vecteur. C'est à la justification détaillée de cette assimilation que nous allons procéder maintenant.

Comme dans le chapitre précédent (on se reportera en particulier à la figure 3.1), on considère entre a et b la trajectoire effectivement suivie (trajectoire de référence) et une trajectoire variée de façon infinitésimale. Les deux trajectoires sont paramétrées par le temps propre de la trajectoire de référence et on se reportera au chapitre précédent pour une description des éléments différentiels importants.

La variation de l'action de la particule libre s'obtient comme au chapitre précédent. En écrivant la variation de l'intervalle élémentaire et en se livrant à la traditionnelle intégration par parties, on trouve:

$$\delta S_{\text{Particule Libre}} = \int_a^b \frac{dP_\mu}{d\tau} \delta x^\mu d\tau \ . \tag{4.5}$$

Calculons maintenant la variation de l'action d'interaction:

$$\begin{aligned} \delta S_{\text{Interaction}} &= -q\delta \int_a^b \mathcal{A}_\mu dx^\mu \\ &= -q \int_a^b [\delta \mathcal{A}_\mu]\, dx^\mu \ - q \int_a^b \mathcal{A}_\mu \delta dx^\mu \ , \end{aligned} \tag{4.6}$$

où le premier terme représente la variation due au fait que la trajectoire variée n'échantillonne pas le potentiel aux mêmes points que la trajectoire de référence. Le deuxième terme représente pour sa part la variation de l'action due à la modification de géométrie de la trajectoire à potentiel constant.

Traitons d'abord le deuxième terme. On a en fait

$$\begin{aligned} -q \int_a^b \mathcal{A}_\mu \delta dx^\mu &= -q \int_a^b \mathcal{A}_\mu \frac{d\delta x^\mu}{d\tau} d\tau \\ &= [-q\mathcal{A}_\mu \delta x^\mu]_a^b + q \int_a^b \left(\frac{d\mathcal{A}_\mu}{d\tau} d\tau \right) \delta x^\mu \ . \end{aligned} \tag{4.7}$$

Le terme tout intégré de l'intégration par parties est identiquement nul, puisque les trajectoires coïncident aux extrémités. La parenthèse dans l'intégrale restante représente l'accroissement $d\mathcal{A}_\mu$ du potentiel quand on passe d'une extrémité à l'autre d'un élément différentiel de la trajectoire de référence. Le potentiel étant une fonction de l'événement auquel il est estimé, nous pourrons écrire:

$$d\mathcal{A}_\mu = \partial_\nu \mathcal{A}_\mu dx^\nu \ . \tag{4.8}$$

Le deuxième terme s'écrit donc:

$$q \int_a^b \partial_\nu \mathcal{A}_\mu dx^\nu \delta x^\mu \ . \tag{4.9}$$

Revenons maintenant au premier terme. La variation $\delta \mathcal{A}_\mu$ du potentiel quand on passe de la trajectoire de référence à la trajectoire variée s'écrit $\delta \mathcal{A}_\mu = \partial_\nu \mathcal{A}_\mu \delta x^\nu$. Le terme a intégrer, après une permutation sans conséquences des indices muets, s'écrit donc $\partial_\mu \mathcal{A}_\nu \delta x^\mu dx^\nu$. La variation de l'action d'interaction peut donc finalement s'écrire:

$$\delta S_{\text{Interaction}} = q \int_a^b [\partial_\nu \mathcal{A}_\mu - \partial_\mu \mathcal{A}_\nu]\, \delta x^\mu dx^\nu \ . \tag{4.10}$$

En remarquant finalement que, sur la trajectoire de référence, $dx^\nu = (dx^\nu/d\tau)d\tau = U^\nu d\tau$, où U^ν est la 4–vitesse, on peut écrire la variation totale de l'action sous la forme:

$$\delta S = \int_a^b \left[\frac{dP_\mu}{d\tau} - G_\mu \right] \delta x^\mu d\tau \ , \tag{4.11}$$

avec

$$G_\mu = q\left[\partial_\mu \mathcal{A}_\nu - \partial_\nu \mathcal{A}_\mu\right] U^\nu \ . \tag{4.12}$$

Les équations de Lagrange s'obtiennent alors immédiatement. La variation de l'action ne peut s'annuler au premier ordre dans les écarts entre les trajectoires que si tous les coefficients des δx^μ sont identiquement nuls. Les équations de mouvement s'écrivent donc:

$$\frac{dP_\mu}{d\tau} = G_\mu \ , \tag{4.13}$$

la quantité G_μ n'étant autre que la 4–force que nous avions pressentie au chapitre précédent. Notons que cette équation 4–vectorielle contient aussi bien la variation de l'énergie de la particule que celle de sa quantité de mouvement.

La force s'exprime en fonction du rotationnel du potentiel, qui est un tenseur de rang deux, antisymétrique, écrit ici sous sa forme complètement covariante. Nous appellerons tenseur champ ce rotationnel et nous poserons:

$$F_{\mu\nu} = \partial_\mu \mathcal{A}_\nu - \partial_\nu \mathcal{A}_\mu \ . \tag{4.14}$$

La force à laquelle est soumise la particule s'écrit alors simplement:

$$G_\mu = qF_{\mu\nu}U^\nu \tag{4.15}$$

et n'est pas autre chose que la contraction du tenseur champ avec la vitesse de la particule.

4.1.2 Tenseur champ électromagnétique

$F_{\mu\nu}$ est par définition un tenseur antisymétrique de rang 2, le 4–rotationnel du potentiel $(V/c, \mathbf{A})$. Il ne dépend donc que de six coordonnées indépendantes. Les trois coordonnées spatio–temporelles forment les composantes d'un vecteur spatial, alors que les trois coordonnées purement spatiales forment les composantes d'un pseudo–vecteur.

On peut écrire les composantes spatio–temporelles sous la forme:

$$\begin{aligned} F_{0i} &= \partial_0 \mathcal{A}_i - \partial_i \mathcal{A}_0 \\ &= \frac{1}{c}\left(-\frac{\partial A_i}{\partial t} - \partial_i V\right) \\ &= \frac{E_i}{c} \ , \end{aligned} \tag{4.16}$$

en posant

$$\mathbf{E} = -\boldsymbol{\nabla} V - \frac{\partial \mathbf{A}}{\partial t} \ . \tag{4.17}$$

Nous nommerons bien sûr "champ électrique" le vrai vecteur spatial ainsi défini.

Nous pouvons de même mettre les composantes purement spatiales du tenseur champ sous la forme:

$$F_{12} = -B_z \tag{4.18}$$

$$F_{13} = B_y \tag{4.19}$$

$$F_{23} = -B_x \ , \tag{4.20}$$

en introduisant le pseudo–vecteur "champ magnétique"

$$\mathbf{B} = \boldsymbol{\nabla} \times \mathbf{A} \tag{4.21}$$

De manière toute naturelle, notre théorie de champ décrite par un 4–vecteur potentiel s'exprime en fonction de deux champs et ressemble de plus en plus à l'électromagnétisme de Maxwell.

On peut écrire F sous forme matricielle:

$$F_{\mu\nu} = \begin{pmatrix} 0 & E_x/c & E_y/c & E_z/c \\ -E_x/c & 0 & -B_z & B_y \\ -E_y/c & B_z & 0 & -B_x \\ -E_z/c & -B_y & B_x & 0 \end{pmatrix} , \tag{4.22}$$

ou encore, sous forme complètement contravariante:

$$F^{\mu\nu} = \begin{pmatrix} 0 & -E_x/c & -E_y/c & -E_z/c \\ E_x/c & 0 & -B_z & B_y \\ E_y/c & B_z & 0 & -B_x \\ E_z/c & -B_y & B_x & 0 \end{pmatrix} . \tag{4.23}$$

L'invariance de jauge est contenue dans la définition même du tenseur champ électromagnétique en fonction du potentiel. L'équation du mouvement de la particule (qui a seule un sens physique non ambigu) est inchangée si nous ajoutons au 4–potentiel un 4–gradient arbitraire (il est aisé de vérifier que le 4–rotationnel d'un champ de 4–gradient est identiquement nul). La transformation "de jauge"

$$\mathcal{A}_\mu \longrightarrow \mathcal{A}_\mu + \partial_\mu \Phi , \tag{4.24}$$

où Φ est un champ de 4–scalaires arbitraire, laisse invariante la dynamique. En exprimant cette transformation en termes des composantes spatiales et temporelles du potentiel, nous retrouvons la forme standard de la transformation de jauge:

$$\begin{aligned} \mathbf{A} &\longrightarrow \mathbf{A} - \boldsymbol{\nabla}\Phi \\ V &\longrightarrow V + \frac{\partial \Phi}{\partial t} . \end{aligned} \tag{4.25}$$

Pour lever l'ambiguïté sur le potentiel, nous pouvons imposer une condition de jauge supplémentaire. Pour respecter l'invariance relativiste, cette condition de jauge se doit d'être manifestement covariante. La plus naturelle, la "jauge de Lorentz", est d'imposer la nullité de la 4–divergence du potentiel:

$$\partial_\mu \mathcal{A}^\mu = 0 , \tag{4.26}$$

qui s'écrit en termes des composantes spatiales et temporelles:

$$\frac{1}{c^2}\frac{dV}{dt} + \boldsymbol{\nabla} \cdot \mathbf{A} = 0 . \tag{4.27}$$

Notons que la jauge de l'électrostatique, ou jauge de Coulomb, $\boldsymbol{\nabla} \cdot \mathbf{A} = 0$, brise la covariance. Si elle peut être employée sans restriction dans un référentiel donné, elle est à proscrire quand s'imposent des changements de référentiel.

4.1.3 Force de Lorentz

Pour nous rapprocher encore de l'électromagnétisme sous sa forme classique, nous allons exprimer la force tridimensionnelle subie par la particule chargée dans un référentiel donné en fonction des champs électriques et magnétiques. En fonction du temps t dans le référentiel $\mathcal{R}$, l'équation de la dynamique s'écrit:

$$\gamma \frac{dP_\mu}{dt} = \gamma(\frac{d\mathcal{E}/c}{dt}, -\frac{d\mathbf{p}}{dt}) = qF_{\mu\nu}U^\nu , \tag{4.28}$$

où $\mathcal{E}$ est l'énergie totale de la particule et $\mathbf{p}$ sa quantité de mouvement tridimensionnelle. En développant le dernier terme (qui peut s'écrire simplement comme un produit matriciel) et en isolant les composantes temporelles et spatiales, on trouve:

$$\frac{d\mathcal{E}}{dt} = q\mathbf{E} \cdot \mathbf{v} , \tag{4.29}$$

où $\mathbf{v}$ est la vitesse de la particule dans $\mathcal{R}$ et

$$\frac{d\mathbf{p}}{dt} = q(\mathbf{E} + \mathbf{v} \times \mathbf{B}) \ . \tag{4.30}$$

On retrouve ainsi, dans un référentiel donné, la forme standard de la force de Lorentz. On trouve aussi que la variation d'énergie est entièrement due au champ électrique. Bien sûr, dans ces équations, $\mathcal{E}$ et $\mathbf{p}$ sont des quantités relativistes ($\mathcal{E} = m\gamma c^2$ et $\mathbf{p} = m\gamma\mathbf{v}$). La dynamique de la particule est donc en général différente de la dynamique classique.

Comme la force de Lorentz est la première force que nous ayons explicitée dans le cadre relativiste, nous allons, à titre d'application et d'exemple, étudier en détail le mouvement d'une particule chargée dans un champ magnétique ou électrique uniforme, statique. Nous verrons ainsi comment la nature relativiste du mouvement modifie la dynamique et nous pourrons jeter un regard nouveau sur la notion de vitesse limite.

Champ magnétique uniforme

Nous nous placerons, dans ce paragraphe et le suivant, dans un référentiel particulier $\mathcal{R}$ et nous abandonnerons donc la covariance manifeste. Nous considérerons le mouvement d'une particule dans un champ magnétique $\mathbf{B}$ uniforme et constant.

Le champ magnétique ne modifiant pas l'énergie totale de la particule, le facteur γ est une constante et l'équation du mouvement s'écrit simplement:

$$\gamma m \frac{d\mathbf{v}}{dt} = q\mathbf{v} \times \mathbf{B} \ . \tag{4.31}$$

Elle est donc la même que dans le cas non relativiste, avec la simple substitution de la masse m par γm. En particulier, la trajectoire est une hélice admettant le champ magnétique pour axe, avec un rayon $R = m\gamma v/qB$ où v est le module de la vitesse perpendiculaire au champ magnétique. la pulsation du mouvement circulaire uniforme dans le plan perpendiculaire à $\mathbf{B}$ (pulsation cyclotron) étant:

$$\omega_c = \frac{qB}{m\gamma} \ . \tag{4.32}$$

S'il n'y a pas de différence qualitative entre le mouvement relativiste et le mouvement classique, le facteur γ induit néanmoins des complications techniques dans les applications. Dans de nombreux types d'accélérateurs, un champ magnétique est utilisé pour confiner les particules au voisinage d'une trajectoire circulaire. Le facteur γ fait que le rayon de ces trajectoires est, pour des particules ultra–relativistes, beaucoup plus grand que ce que prédit la mécanique classique. L'encombrement de ce type de dispositif est en partie dû à cet effet[1]. De plus, la fréquence des champs accélérateurs, qui doit être adaptée à la fréquence cyclotron, doit être ajustée pendant toute la phase d'accélération pour tenir compte de la variation de ce facteur relativiste.

Notons que cette "contraction" relativiste de la fréquence cyclotron peut être mise en évidence même pour des électrons de très basse énergie. Dans une très spectaculaire série d'expériences, Hans Dehmelt et ses collaborateurs (Université de Seattle) ont étudié des électrons confinés dans un piège constitué d'un champ magnétique et d'un champ quadripolaire électrique (piège de Penning). Ils ont ainsi mesuré avec une précision remarquable, sur un électron unique, le célèbre "facteur gyromagnétique anormal", qui constitue un test sévère de l'électrodynamique quantique. Une des étapes de l'expérience est d'exciter, par un champ radiofréquence convenable, le mouvement cyclotron de l'électron. Dehmelt a pu observer que la fréquence de résonance cyclotron se déplace avec l'énergie de l'électron, conformément à la loi relativiste. Les énergies mises en jeu n'étant que d'une fraction d'électron–volt, on pourra juger de la sensibilité de l'expérience.

[1] Un autre effet important limite la compacité des accélérateurs: plus une particule est accélérée, plus elle perd d'énergie par rayonnement. Nous discuterons de ce "rayonnement de freinage" dans un prochain chapitre. Pour des particules légères (électrons), cet effet est la principale limitation au rayon des accélérateurs. Notons aussi que les accélérateurs linéaires échappent à ces deux types de limitations.

Champ électrique uniforme

Nous considérerons maintenant le mouvement d'une particule dans un champ électrique uniforme et constant. Pour simplifier l'algèbre, sans trop restreindre la physique, nous prendrons comme condition initiale une particule au repos. Le champ électrique étant par convention orienté le long de l'axe Ox, il est évident que le mouvement s'effectue le long de cet axe. L'équation de la dynamique, projetée sur Ox, s'écrit alors:

$$m\frac{d\gamma\dot{x}}{dt} = qE \ . \tag{4.33}$$

Dans ce cas, bien sûr, l'énergie de la particule et donc le facteur γ ne sont pas des constantes. On tire de cette équation immédiatement:

$$\gamma\dot{x} = \frac{qEt}{m} \ , \tag{4.34}$$

la valeur initiale de cette quantité étant nulle par convention. Nous poserons, pour alléger les notations, $V = qEt/m$ (notons que V serait la vitesse de la particule si nous ne tenions pas compte des corrections relativistes au mouvement). On déduit alors de ce qui précède:

$$\dot{x} = \frac{V}{\sqrt{1+V^2/c^2}} \ , \tag{4.35}$$

qui s'intégre aisément en

$$x = \frac{mc^2}{qE}\left[\sqrt{1+\left(\frac{qEt}{mc}\right)^2} - 1\right] \ , \tag{4.36}$$

à condition de prendre $x = 0$ comme condition initiale.

Pour des temps suffisamment petits, la vitesse de la particule est faible et on peut développer l'expression précédente au premier ordre en qEt/mc. On trouve alors

$$x = \frac{1}{2m}qEt^2 \ , \tag{4.37}$$

mouvement uniformément accéléré de la dynamique classique. Aux temps longs, en revanche, x tend simplement vers ct: la vitesse de la particule tend vers la vitesse de la lumière, comme nous pouvions nous y attendre. On pourra montrer, à titre d'exercice, que la rapidité de la particule continue, pour sa part, à croître indéfiniment. La généralisation de ce calcul à trois dimensions ne pose aucune autre difficulté qu'algébrique.

4.1.4 Changements de référentiels pour le champ

La formulation explicitement relativiste du tenseur champ électromagnétique nous permet d'écrire sans difficultés la loi de transformation des champs dans un changement de référentiel. Nous aurons en effet:

$$F'^{\mu\nu}(x'^{\alpha} = \mathcal{L}^{\alpha}{}_{\beta}x^{\beta}) = \mathcal{L}^{\mu}{}_{\rho}\mathcal{L}^{\nu}{}_{\sigma}F^{\rho\sigma}(x^{\beta}) \ , \tag{4.38}$$

où les quantités primées sont relatives au nouveau référentiel. Le champ étant un champ de tenseur, il est une fonction de l'événement auquel il est estimé. Il faut bien sûr estimer le champ dans les deux référentiels pris au même événement et donc à des coordonnées spatio–temporelles qui se déduisent les unes des autres dans une transformation de Lorentz. Dans l'expression précédente, $\mathcal{L}$ peut décrire n'importe quel élément du groupe de Lorentz le plus général.

Nous préciserons maintenant les nouvelles valeurs du champ pour une transformation spéciale de Lorentz avec les conventions habituelles pour l'orientation des axes. Le calcul ne présente aucune difficulté de principe. Il faut toutefois prendre garde que le produit de "tenseurs" au second membre ne peut être calculé directement comme un produit de leurs trois représentations matricielles. Les deux

derniers termes correspondent en effet à une sommation sur deux indices colonne. On peut mettre le calcul sous la forme de produits matriciels standard en effectuant les transpositions nécessaires. Après quelques lignes d'algèbre, on arrive aux lois suivantes pour les transformations des composantes des champs électriques et magnétiques:

$$E'_x = E_x \tag{4.39}$$
$$E'_y = \gamma(E_y - c\beta B_z) \tag{4.40}$$
$$E'_z = \gamma(E_z + c\beta B_y) \tag{4.41}$$
$$B'_x = B_x \tag{4.42}$$
$$B'_y = \gamma(B_y + \beta\frac{E_z}{c}) \tag{4.43}$$
$$B'_z = \gamma(B_z - \beta\frac{E_y}{c}) \ . \tag{4.44}$$

La transformation inverse s'obtient trivialement en changeant le signe de β[2]. Si, jusqu'alors, nous n'avions fait que retrouver les caractéristiques essentielles de l'électromagnétisme (il n'est peut-être plus utile de cacher que notre théorie de champ est bien l'électromagnétisme), nous obtenons ici, grâce à la formulation manifestement covariante, un résultat nouveau et fort important.

On peut obtenir une approximation galiléenne à la loi de transformation des champs en ne gardant que l'ordre le plus bas en u/c dans les équations précédentes. On obtient alors, sous forme vectorielle:

$$\mathbf{E}' = \mathbf{E} - \mathbf{B} \times \mathbf{u} \tag{4.45}$$
$$\mathbf{B}' = \mathbf{B} + \frac{\mathbf{E} \times \mathbf{u}}{c^2} \ . \tag{4.46}$$

4.1.5 Invariants du champ électromagnétique

On peut se poser le problème de former des quantités 4–scalaires à partir du tenseur champ. De telles quantités seront en effet conservées dans un changement de référentiel et constitueront des invariants du champ électromagnétique, fort utiles. Là encore, nous allons ajouter des résultats nouveaux à l'électromagnétisme standard. Nous ne chercherons pas systématiquement tous les invariants possibles. En fait, il n'en existe que deux qui présentent un intérêt physique[3].

Formons d'abord la quantité:

$$F_{\mu\nu}F^{\mu\nu} \ . \tag{4.47}$$

Cette quantité est manifestement un 4–scalaire et donc un invariant du champ. Ecrivons–la en termes des champs électriques et magnétiques pour en comprendre la signification physique. Interviennent des composantes spatio–temporelles et des composantes spatiales. La contribution des composantes spatio–temporelles est manifestement $2F_{0i}F^{0i}$ (en effet les deux termes se déduisant l'un de l'autre par permutation des indices sont manifestement égaux en raison de l'antisymétrie de F) ou encore $-2E^2/c^2$. De même, les composantes spatiales font intervenir le carré scalaire du champ magnétique et on a enfin:

$$F_{\mu\nu}F^{\mu\nu} = \frac{2}{c^2}(c^2B^2 - E^2) \ . \tag{4.48}$$

La quantité $c^2B^2 - E^2$ est donc un invariant du champ.

Donnons dès maintenant une application de cet invariant. Dans un référentiel $\mathcal{R}$, considérons une onde électromagnétique plane. Les modules du champ électrique et du champ magnétique sont reliés

[2]On notera que ces expressions ne sont pas invariantes par échange de y et z. Ceci n'est en rien contradictoire avec la symétrie du problème. Echanger ces axes revient à changer l'orientation de l'espace et donc le signe de $\mathbf{B}$ qui est un pseudo-vecteur.

[3]On montrera en particulier que le déterminant de $F^{\mu\nu}$ est proportionnel au carré de notre second invariant

par $E = cB$. L'invariant du champ considéré est donc nul. Dans un autre référentiel, sans préjuger de la structure du champ[4], on aura aussi $E' = cB'$.

Le deuxième invariant que nous allons former s'écrit:

$$\epsilon^{\mu\nu\rho\sigma} F_{\mu\nu} F_{\rho\sigma} \ , \tag{4.49}$$

où ϵ est le tenseur complètement antisymétrique de rang 4. Rappelons rapidement que les éléments de ce tenseur valent +1 si les quatre indices sont une permutation paire de $(0, 1, 2, 3)$, -1 si ils constituent une permutation impaire, 0 dans tous les autre cas.

Il n'y a donc dans cet invariant que 24 termes non nuls. En fait, ces termes sont égaux 8 à 8. Si nous considérons une permutation μ, ν, ρ, σ donnée, nous obtenons en effet un terme identique en échangeant μ et ν et/ou ρ et σ (le tenseur antisymétrique et le tenseur champ concernés changent tous les deux de signe). Il y a donc au total quatre permutations de ce type et 4 termes identiques dans le développement de notre invariant. De plus, nous obtenons un terme identique en échangeant la première paire (μ, ν) et la seconde (ρ, σ). En effet les termes en tenseur champ ne changent pas. Cette permutation des deux paires peut s'effectuer avec quatre permutations des indices. Elle ne change pas non plus la valeur du tenseur antisymétrique. Ensuite, les opérations de permutations à l'intérieur des paires ainsi permutées peuvent être effectuées sans changer la valeur. Nous introduisons ainsi quatre nouvelles permutations des indices donnant la même valeur. Au total, il y a donc 8 termes identiques. Comme nous n'avons manifestement que 24 termes non nuls, il n'y a que trois termes différents, correspondant à un ensemble d'indices (par exemple 0,1,2,3) et aux deux manières d'échanger un terme de la première paire et un terme de la deuxième (dans ce cas 2,1,0,3 et 3,1,2,0). Le calcul de ces trois termes est alors trivial à partir de l'expression de F (ϵ valant +1 pour la première permutation et -1 pour les deux autres). On trouve finalement:

$$\epsilon^{\mu\nu\rho\sigma} F_{\mu\nu} F_{\rho\sigma} = -8\mathbf{E} \cdot \mathbf{B}/c \ , \tag{4.50}$$

un résultat particulièrement simple.

Le produit scalaire des champs électriques et magnétiques est donc invariant dans un changement de référentiel (bien sûr, cette invariance pourrait être établie, de manière assez pénible, directement à partir des lois de transformation). Donnons tout de suite une application de cette propriété. Dans un référentiel, considérons une onde plane. **E** et **B** sont alors perpendiculaires et leurs modules sont dans un rapport c. Dans un autre référentiel, ils sont donc encore perpendiculaires avec des modules dans un rapport c.

4.1.6 Premier groupe d'équations de Maxwell

Il nous reste à tirer parti du fait que le tenseur champ est le rotationnel du potentiel. A trois dimensions, cette propriété impliquerait la nullité de sa divergence. Nous allons maintenant établir la propriété correspondante à quatre dimensions. En écrivant cette propriété en termes des champs électriques et magnétiques, nous établirons des relations différentielles entre eux qui ne seront autres que les deux équations de Maxwell homogènes. Nous avons donc

$$F_{\mu\nu} = \partial_\mu \mathcal{A}_\nu - \partial_\nu \mathcal{A}_\mu \ . \tag{4.51}$$

On en déduit immédiatement:

$$\begin{aligned} \partial_\rho F_{\mu\nu} &= \partial_\rho \partial_\mu \mathcal{A}_\nu - \partial_\rho \partial_\nu \mathcal{A}_\mu \\ \partial_\nu F_{\rho\mu} &= \partial_\nu \partial_\rho \mathcal{A}_\mu - \partial_\nu \partial_\mu \mathcal{A}_\rho \\ \partial_\mu F_{\nu\rho} &= \partial_\mu \partial_\nu \mathcal{A}_\rho - \partial_\mu \partial_\rho \mathcal{A}_\nu \end{aligned} \tag{4.52}$$

[4]Qui se trouve être également une onde plane.

En remarquant que les dérivées secondes croisées du potentiel sont égales, et en faisant la somme de ces trois équations, on obtient immédiatement:

$$\partial_\rho F_{\mu\nu} + \partial_\nu F_{\rho\mu} + \partial_\mu F_{\nu\rho} = 0 \ . \tag{4.53}$$

Cette équation est une conséquence directe du fait que le champ dérive d'un potentiel (c'est d'ailleurs une condition nécessaire et suffisante).

Pour en comprendre la signification physique, écrivons cette équation en termes des champs électriques et magnétiques. Remarquons d'abord que, s'il y a a priori 64 équations possibles, seules 4 ne sont pas triviales. Si les trois indices sont identiques, tous les F sont nuls et l'équation est un truisme. Si deux indices sont égaux (par exemple $\rho = \mu$), l'équation se réduit à $\partial_\mu(F_{\mu\nu} + F_{\nu\mu}) = 0$, une tautologie en raison de l'antisymétrie de F. L'équation n'est non triviale que si les trois indices sont différents et il ne reste donc que quatre équations indépendantes.

La première correspond aux indices 1,2,3. Elle s'écrit:

$$\partial_1 F_{23} + \partial_3 F_{12} + \partial_2 F_{31} = 0 \ , \tag{4.54}$$

soit encore

$$\boldsymbol{\nabla} \cdot \mathbf{B} = 0 \ . \tag{4.55}$$

On montrera de même que les trois autres équations peuvent se résumer, sous forme vectorielle, par:

$$\boldsymbol{\nabla} \times \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t} \ . \tag{4.56}$$

Nous retrouvons donc ainsi les équations de Maxwell homogènes, qui sont équivalentes à l'existence d'un potentiel scalaire et d'un potentiel vecteur.

4.2 Champ en fonction des sources

Nous allons maintenant établir les équations qui relient le tenseur champ à ses sources, c'est à dire au mouvement des particules chargées. Dans le paragraphe précédent, nous nous intéressions au mouvement d'une particule unique. Les variables dynamiques du problème étaient donc la position et l'impulsion de la particule, situation habituelle en mécanique analytique. Dans tout ce chapitre, conformément à notre programme initial, nous supposerons imposées les dynamiques des particules (c'est-à-dire le courant) et nous ne nous intéresserons qu'à la dynamique du champ. Les variables dynamiques sont donc les valeurs du potentiel ou des champs en tous points de l'espace et à chaque instant. Il nous faudra donc adapter nos techniques variationnelles pour des variables dynamiques continues. En particulier, nous n'écrirons plus l'action en termes de lagrangien mais d'une densité de lagrangien que nous intégrerons sur tout l'espace et sur le temps pour obtenir l'action. Il nous faudra aussi réécrire l'action d'interaction comme l'intégrale sur tout l'espace d'une densité de lagrangien qui devra faire intervenir le 4–vecteur courant au lieu des positions et vitesses individuelles des particules.

Le fait que nous traitions de plusieurs particules pose une difficulté technique immédiate. Pour obtenir les équations du mouvement d'une particule unique, nous avons intégré le lagrangien entre deux événements limites relatifs à cette particule. Nous ne pouvons définir de façon aussi simple les bornes d'intégration si nous considérons plusieurs particules qui ne partagent pas le même temps propre. Pour éviter toute difficulté ou le recours à un formalisme complexe nous éluderons le problème en nous plaçant, pour un temps, dans un référentiel donné $\mathcal{R}$. Dans ce référentiel, le temps est bien défini, et nous pourrons intégrer la densité de lagrangien entre deux instants de référence. Nous n'aurons à manipuler que des intégrales d'espace et de temps sous forme habituelle. En abandonnant ainsi la covariance manifeste, nous risquons bien sûr d'obtenir des équations de Lagrange qui ne seraient pas des invariants relativistes. Nous verrons qu'il n'en sera heureusement rien: les équations que nous obtiendrons dans un référentiel donné s'écriront en termes de quantités explicitement covariantes, et seront donc valables dans n'importe quel référentiel.

4.2.1 Interaction champ–courant

Nous allons, dans ce premier paragraphe, réécrire l'action d'interaction en termes du courant macroscopique. Nous considérons un ensemble de particules chargées ponctuelles, dont le mouvement est imposé, que nous indicerons par un indice α entre parenthèses, pour éviter toute confusion entre cet indice qui numérote simplement les particules et un indice relativiste en position covariante.

Le 4–vecteur courant J^μ peut donc s'écrire: $J^\mu = (c\rho, \mathbf{j})$, avec:

$$\rho = \sum_\alpha q_{(\alpha)} \delta(\mathbf{r} - \mathbf{r}_{(\alpha)}) \tag{4.57}$$

$$\mathbf{j} = \sum_\alpha q_{(\alpha)} \mathbf{v}_{(\alpha)} \delta(\mathbf{r} - \mathbf{r}_{(\alpha)}) \ , \tag{4.58}$$

où $q_{(\alpha)}, \mathbf{r}_{(\alpha)}$ et $\mathbf{v}_{(\alpha)}$ sont respectivement la charge, la position et la vitesse de la particule α. En généralisant l'action d'interaction introduite au paragraphe précédent, on écrira:

$$S_{\text{Interaction}} = -\sum_\alpha q_{(\alpha)} \int_{t_a}^{t_b} \mathcal{A}_\mu({x^\nu}_{(\alpha)}) U^\mu d\tau_{(\alpha)} \ , \tag{4.59}$$

où ${x^\nu}_{(\alpha)}$ est la position de la particule α. t_a et t_b sont les instants dans $\mathcal{R}$ où nous spécifierons les conditions initiales imposées au champ. $\tau_{(\alpha)}$ est le temps propre de la particule α, qui peut être paramétré lui même par le temps t du référentiel $\mathcal{R}$ dans lequel nous nous sommes placés. En raison de la "dilatation des temps", $d\tau_{(\alpha)} = dt/\gamma_{(\alpha)}$ où $\gamma_{(\alpha)}$ est le facteur de dilatation relativiste calculé à chaque instant avec la vitesse de la particule α. En substituant cette expression dans l'action d'interaction et en explicitant enfin les composantes temporelles et spatiales, on trouve:

$$S_{\text{Interaction}} = -\int_{t_a}^{t_b} dt \ \left[\sum_\alpha \mathcal{A}_\mu({x_{(\alpha)}}^\nu)(q_{(\alpha)}c, q_{(\alpha)}\mathbf{v}_{(\alpha)}) \right] \ . \tag{4.60}$$

Notons que l'équation précédente, qui mélange notations d'Einstein et séparation des parties temporelles et spatiales constitue un abus de notations manifeste. Dans cette expression, le potentiel est à évaluer à l'endroit où se trouve la particule α. Pour mettre l'expression précédente sous la forme de l'intégrale d'une densité de lagrangien, on peut écrire:

$$\mathcal{A}_\mu({x_{(\alpha)}}^\nu) = \int dV \ \mathcal{A}_\mu(ct, \mathbf{r}) \delta(\mathbf{r} - \mathbf{r}_{(\alpha(t))}) \ , \tag{4.61}$$

l'intégrale portant sur tout l'espace. On a alors:

$$S_{\text{Interaction}} = -\int_{t_a}^{t_b} dt \int dV \ \left[\sum_\alpha (q_{(\alpha)}c, q_{(\alpha)}\mathbf{v}_{(\alpha)}) \delta(\mathbf{r} - \mathbf{r}_{(\alpha)}(t)) \right] \mathcal{A}_\mu(ct, \mathbf{r}) \tag{4.62}$$

On reconnaît, entre les crochets, l'expression du 4–vecteur courant. On a donc finalement:

$$S_{\text{Interaction}} = \int_{t_a}^{t_b} dt \int dV \ \mathcal{L}_{\text{Interaction}} \tag{4.63}$$

où la densité de lagrangien d'interaction s'exprime par:

$$\mathcal{L}_{\text{Interaction}} = -\mathcal{A}_\mu J^\mu \ . \tag{4.64}$$

Bien que nous ayons établi cette expression dans un référentiel donné, elle est manifestement un 4–scalaire et est donc correcte dans tous les référentiels.

4.2.2 Lagrangien du champ

Il nous faut maintenant postuler l'expression de la densité de lagrangien pour le champ libre. Nous allons choisir bien sûr une quantité qui soit manifestement un 4–scalaire. Il faudra de plus qu'elle soit invariante de jauge et donc qu'elle ne s'exprime en définitive qu'en fonction du tenseur champ et non du potentiel. Enfin, il faudra que ce soit une quantité quadratique dans le champ. Une densité de lagrangien est en effet homogène à une densité d'énergie qui doit être une fonction quadratique des variables dynamiques. Les deux invariants scalaires du champ que nous avions construits au paragraphe précédent remplissent tous deux ces conditions. Le second, équivalent au produit scalaire des champs électriques et magnétiques, ne semble pas convenir. On ne voit pas, en particulier, comment il pourrait décrire de façon convenable une situation purement électrostatique. Seule l'autre invariant est convenable et nous postulerons que la densité de lagrangien pour le champ libre peut s'écrire:

$$\mathcal{L}_{\text{Champ Libre}} = -\frac{1}{4\mu_0} F_{\mu\nu} F^{\mu\nu} \ , \tag{4.65}$$

où μ_0 est a priori une constante dimensionnelle telle que la densité de lagrangien ait la dimension d'une densité d'énergie. L'action totale intégrée entre les instants t_a et t_b s'écrit donc:

$$S = S_{\text{Particules Libres}} + \int_{t_a}^{t_b} dt \int dV \ \left[-\mathcal{A}_\mu J^\mu - \frac{1}{4\mu_0} F_{\mu\nu} F^{\mu\nu} \right] \ . \tag{4.66}$$

Dans cette expression, l'action des particules libres est une simple constante, puisque leur dynamique est imposée.

4.2.3 Equations de Lagrange

Pour établir les équations de Lagrange, nous allons considérer, entre les deux instants de référence t_a et t_b, deux histoires possibles du champ. L'une, qui sera la "trajectoire effectivement suivie", autrement dit la solution des équations de Lagrange, correspondra, à chaque instant, au potentiel $\mathcal{A}_\mu$. L'autre, infinitésimalement différente, correspondra au potentiel $\mathcal{A}_\mu + \delta\mathcal{A}_\mu$. Pour assurer que le champ vrai et le champ varié obéissent aux mêmes conditions aux limites, nous imposerons à l'accroissement infinitésimal $\delta\mathcal{A}$ de s'annuler, en tous points de l'espace, en t_a et en t_b. Nous allons ensuite exprimer la variation de l'action due à cette variation du potentiel en tous points de l'espace à chaque instant. En exprimant que cette variation est nulle au premier ordre dans l'accroissement, quel que soit celui-ci, nous obtiendrons une relation qui devra être vérifiée par le champ en tous points, à tout instant. Ce raisonnement généralise de façon évidente à un ensemble continu de variables dynamiques celui que nous avons déjà utilisé fréquemment pour un nombre fini de degrés de liberté.

La trajectoire des particules étant imposée, le courant ne doit pas être varié et la variation de l'action totale, s'écrit donc:

$$\delta S = -\int_{t_a}^{t_b} dt \int dV \ \left[\delta(\mathcal{A}_\mu) J^\mu + \frac{1}{4\mu_0} \delta(F_{\mu\nu} F^{\mu\nu}) \right] \ . \tag{4.67}$$

On a de manière évidente:

$$\delta(F_{\mu\nu} F^{\mu\nu}) = 2(\delta F_{\mu\nu}) F^{\mu\nu} \ . \tag{4.68}$$

En exprimant ensuite $\delta F_{\mu\nu}$ en termes du potentiel, on a:

$$\delta S = -\int_{t_a}^{t_b} dt \int dV \ \left[\delta(\mathcal{A}_\mu) J^\mu + \frac{1}{2\mu_0} F^{\mu\nu} \partial_\mu \delta\mathcal{A}_\nu - \frac{1}{2\mu_0} F^{\mu\nu} \partial_\nu \delta\mathcal{A}_\mu \right] \ . \tag{4.69}$$

En permutant les deux indices muets du terme central, et en utilisant l'antisymétrie de F, on constate que les deux derniers termes dans l'intégrale sont égaux, et que donc:

$$\delta S = -\int_{t_a}^{t_b} dt \int dV \ \left[\delta(\mathcal{A}_\mu) J^\mu - \frac{1}{\mu_0} F^{\mu\nu} \partial_\nu \delta\mathcal{A}_\mu \right] \ . \tag{4.70}$$

Considérons le deuxième terme dans l'intégrale. A un facteur $1/c$ près, il s'agit de l'intégrale d'espace temps d'une quantité scalaire. Le volume d'intégration $\mathcal{V}$ est l'ensemble de l'espace à trois dimensions pris entre les instants t_a et t_b. En utilisant le théorème d'Ostrogradski pour les intégrales quadridimensionnelles, on peut réaliser sur ce terme une intégration par parties. En posant $d\Omega = cdtdV$ on a en effet:

$$\int_{\mathcal{V}} d\Omega\ F^{\mu\nu}\partial_\nu \delta\mathcal{A}_\mu = \int_{\mathcal{S}} F^{\mu\nu}\delta\mathcal{A}_\mu\, dS_\nu - \int_{\mathcal{V}} d\Omega\ (\partial_\nu F^{\mu\nu})\delta\mathcal{A}_\mu\ , \tag{4.71}$$

la surface $\mathcal{S}$, à trois dimensions, étant la frontière du volume $\mathcal{V}$ et dS_ν l'élément différentiel de cette surface. Cette surface est constitué de l'ensemble de l'espace pris à l'instant initial t_a, de la sphère de l'infini (une sphère de rayon R, dont on prend la limite pour $R \to \infty$) à tous les instants entre t_a et t_b et à nouveau de tout l'espace à l'instant final t_b. L'accroissement du potentiel est nul aux instants limites en tous points de l'espace. Il est nul aussi, ainsi que tous les champs physiques, en tous points de la sphère de l'infini à chaque instant. L'intégrale de surface dans l'expression précédente est donc identiquement nulle. Notons que nous retrouvons ici, sous une forme un peu plus complexe, l'intégration par parties à laquelle nous devons toujours procéder pour établir les équations de Lagrange.

L'accroissement de l'action s'écrit donc finalement:

$$\delta S = -\int_{t_a}^{t_b} dt \int dV\ \left[J^\mu + \frac{1}{\mu_0}\partial_\nu F^{\mu\nu}\right]\delta\mathcal{A}_\mu\ . \tag{4.72}$$

Elle ne peut être nulle quel que soit l'accroissement du potentiel que si le champ entre crochets est identiquement nul. Les équations de Lagrange déterminant le champ en fonction des sources s'écrivent donc simplement:

$$\partial_\nu F^{\mu\nu} = -\mu_0 J^\mu\ , \tag{4.73}$$

ou encore

$$\partial_\mu F^{\mu\nu} = \mu_0 J^\nu \tag{4.74}$$

en exploitant les propriétés de symétrie du tenseur champ.

Dans tout ce raisonnement, nous avons abandonné la covariance manifeste en nous plaçant dans un référentiel donné. En revanche, les équations obtenues ne font intervenir que des quantités covariantes. Elles sont donc très générales, et valables dans tout référentiel galiléen.

Nous pouvons maintenant écrire simplement ces équations en termes du potentiel. En reportant l'expression de F, nous avons:

$$\partial_\mu\partial^\mu\mathcal{A}^\nu - \partial_\mu\partial^\nu\mathcal{A}^\mu = \mu_0 J^\nu\ . \tag{4.75}$$

Si nous imposons au potentiel vecteur d'obéir à la condition de Jauge de Lorentz $\partial_\mu\mathcal{A}^\mu = 0$, le deuxième terme s'annule (on permutera les dérivées partielles pour le constater). L'équation aux potentiels s'écrit alors:

$$\partial_\mu\partial^\mu\mathcal{A}^\nu = \mu_0 J^\nu\ . \tag{4.76}$$

L'opérateur différentiel n'est autre que le carré de la norme du gradient: c'est le 4–laplacien ou encore le d'alembertien.

4.2.4 Equations de Maxwell

Pour mettre ces équations sous une forme plus familière, nous allons les exprimer en termes des champs électriques et magnétiques. La partie temporelle de cette équation vectorielle s'écrit en effet:

$$\partial_\nu F^{0\nu} = -\mu_0 c\rho\ , \tag{4.77}$$

ou encore

$$\boldsymbol{\nabla}\cdot\mathbf{E} = \mu_0 c^2\rho = \rho/\epsilon_0\ , \tag{4.78}$$

en posant évidemment $\epsilon_0\mu_0 c^2 = 1$.

Les trois composantes spatiales se mettent de leur côté évidemment sous la forme:

$$\boldsymbol{\nabla} \times \mathbf{B} = \mu_0 \left[\mathbf{j} + \epsilon_0 \frac{\partial \mathbf{E}}{\partial t} \right] . \tag{4.79}$$

Les équations aux potentiels, pourvu que ceux-ci obéissent à la Jauge de Lorentz, se mettent sous la forme:

$$\Box \mathbf{A} = \mu_0 \mathbf{j} \tag{4.80}$$

$$\Box V = \rho/\epsilon_0 , \tag{4.81}$$

où $\Box$ est l'opérateur d'alembertien $(1/c^2)\partial^2/\partial t^2 - \Delta$, Δ étant le laplacien.

Les équations de Lagrange décrivant le champ en fonction des charges sont donc bien les équations de Maxwell faisant intervenir les courants. Si on y ajoute les équations de Maxwell homogènes et les équations de Lagrange pour la dynamique des particules qui sont équivalentes à la force de Lorentz on peut déterminer complètement, au moins en principe, la dynamique couplée des particules et du champ. L'électromagnétisme de Maxwell, formulé en termes explicitement covariants, n'est donc que la théorie de champ la plus simple qui soit décrite par un champ de vecteurs. La structure à deux champs de la théorie de Maxwell est une conséquence immédiate de la nature antisymétrique du tenseur exprimant le rotationnel du potentiel. En fait, l'électromagnétisme de Maxwell, sous sa forme standard, est déjà écrit en termes explicitement relativistes ce qui le rend bien sûr incompatible avec la cinématique classique. La découverte des équations de Maxwell aurait été une conséquence immédiate de celle de la relativité. Mais il fallait bien les difficultés soulevées par l'électromagnétisme de Maxwell pour qu'on songe à mettre en doute la mécanique Newtonienne, parfaitement vérifiée par ailleurs.

4.3 Energie–impulsion du champ

Le champ électromagnétique doit posséder une densité d'énergie. Energie et quantité de mouvement n'étant que deux aspects du même 4–vecteur impulsion, il doit aussi exister une densité de quantité de mouvement pour le champ électromagnétique. Enfin, le champ électromagnétique obéissant, comme les particules matérielles, aux grandes propriétés de symétrie par translation dans l'espace ou dans le temps, ces densités d'énergie et d'impulsion, associées à celles des particules, doivent obéir à des lois de conservation. Nous allons, dans ce paragraphe, définir les densités d'énergie et d'impulsion et établir leurs équations bilan, c'est-à-dire étudier leur propagation. Nous pourrions partir de la formulation lagrangienne de l'électromagnétisme et appliquer les grandes lois de symétrie pour établir la forme de ces densités, comme nous l'avions fait pour une particule libre. Nous nous contenterons en fait de postuler la forme d'un tenseur rassemblant ces quantités, d'établir ses propriétés et d'identifier les différents termes. Cette approche est beaucoup moins satisfaisante qu'une approche à partir des premiers principes, mais elle est beaucoup plus compacte.

Pour ce qui est du bilan d'énergie du champ électromagnétique, nous allons bien sûr retrouver des résultats bien connus sur la densité d'énergie du champ électromagnétique et sa propagation décrite par le vecteur de Poynting, dont le flux décrit le transport d'énergie à travers une surface. Nous ne ferons que rappeler très brièvement les propriétés essentielles de ces quantités. Pour la quantité de mouvement, nous obtiendrons en revanche des résultats nouveaux. Nous établirons la forme d'une densité (bien sûr vectorielle) de quantité de mouvement à trois dimensions. Le transfert d'impulsion à travers une surface s'écrira comme le flux d'une quantité tensorielle de rang 2.

4.3.1 Tenseur énergie–impulsion

Posons

$$\theta^{\alpha\beta} = \frac{1}{\mu_0} \left[g^{\alpha\mu} F_{\mu\lambda} F^{\lambda\beta} + \frac{1}{4} g^{\alpha\beta} F_{\mu\lambda} F^{\mu\lambda} \right] . \tag{4.82}$$

Nous nommerons, pour des raisons qui apparaîtront dans un instant, ce tenseur de rang 2, écrit ici sous sa forme doublement contravariante, le tenseur d'énergie impulsion du champ ("stress tensor" dans la littérature anglo-saxonne). Il s'agit d'un tenseur symétrique. La symétrie du deuxième terme du crochet est manifeste. Pour vérifier celle du premier terme, nous écrirons:

$$\begin{aligned} g^{\beta\mu}F_{\mu\lambda}F^{\lambda\alpha} &= F^{\lambda\alpha}F^{\beta}{}_{\lambda} = F_{\lambda}{}^{\alpha}F^{\beta\lambda} \\ &= F^{\alpha}{}_{\lambda}F^{\lambda\beta} = g^{\alpha\mu}F_{\mu\lambda}F^{\lambda\beta} \ . \end{aligned} \tag{4.83}$$

On a donc bien $\theta^{\alpha\beta} = \theta^{\beta\alpha}$.

Pour nous convaincre de la possible utilité de ce tenseur, nous allons l'écrire en fonction des champs électrique et magnétique. Le deuxième terme est le produit du tenseur métrique, diagonal, par l'invariant $2(B^2 - E^2/c^2)$, densité de lagrangien du champ électromagnétique. Le premier terme fait intervenir $F_{\mu\lambda}F^{\lambda\beta}$ qui se calcule comme un produit matriciel ordinaire. L'action du tenseur métrique est de changer le signe de toutes les lignes ayant un indice spatial dans la représentation matricielle obtenue. En regroupant avec le second terme et après quelques manipulations élémentaires, on peut écrire le tenseur énergie–impulsion sous la forme:

$$\theta^{\alpha\beta} = \begin{pmatrix} u & \Pi_x/c & \Pi_y/c & \Pi_z/c \\ \Pi_x/c & & & \\ \Pi_y/c & & (\boldsymbol{\mathcal{T}}) & \\ \Pi_z/c & & & \end{pmatrix} , \tag{4.84}$$

où

$$u = \frac{\epsilon_0 E^2}{2} + \frac{B^2}{2\mu_0} \tag{4.85}$$

et

$$\mathbf{\Pi} = \frac{\mathbf{E} \times \mathbf{B}}{\mu_0} \ . \tag{4.86}$$

On reconnaîtra bien sûr ici la densité d'énergie électromagnétique et le vecteur de Poynting.

$\boldsymbol{\mathcal{T}}$ est un tenseur purement spatial de rang 2, que nous nommerons "tenseur de Maxwell". L'expression de ses composantes est:

$$\mathcal{T}_{ij} = \epsilon_0 \left[\frac{E^2}{2}\delta_{ij} - E_i E_j \right] + \frac{1}{\mu_0} \left[\frac{B^2}{2}\delta_{ij} - B_i B_j \right] , \tag{4.87}$$

où les δ_{ij} sont simplement les symboles de Kronecker. Notons ici que nous faisons une entorse sérieuse à nos conventions de notations. Quand il nous arrivera de manipuler une quantité, vectorielle ou tensorielle qui soit uniquement relative à l'espace ordinaire à trois dimensions, nous placerons tous les indices en position basse, en appliquant donc avec précautions la règle de sommation sur les indices répétés. Notons enfin que pour un tenseur spatial les composantes complètement contravariantes et complètement covariantes coïncident. Nous allons interpréter plus tard la signification physique de ce tenseur.

4.3.2 Lois de conservation. Interprétation

Si θ est associé à la propagation de l'énergie et de la quantité de mouvement, il doit vérifier des équations locales de conservation qui font intervenir sa divergence (on se souviendra de l'équation locale de bilan d'énergie qui fait intervenir la dérivée temporelle de u et la divergence de $\mathbf{\Pi}$, qui sont rassemblées dans la 4–divergence de la première ligne de θ). Nous allons donc calculer le 4–vecteur

$$\partial_\alpha \theta^{\alpha\beta} \tag{4.88}$$

En remarquant que les tenseurs métriques ne se dérivent pas, et que leur seule action est d'élever l'indice des dérivations, on met ce terme sous la forme:

$$\partial_\alpha \theta^{\alpha\beta} = \frac{1}{\mu_0}\left[\partial^\mu\left(F_{\mu\lambda}F^{\lambda\beta}\right) + \frac{1}{4}\partial^\beta\left(F_{\mu\lambda}F^{\mu\lambda}\right)\right] \ . \tag{4.89}$$

Le dernier terme peut s'écrire:

$$(\partial^\beta F_{\mu\lambda})F^{\mu\lambda} + F_{\mu\lambda}\partial^\beta F^{\mu\lambda} \ . \tag{4.90}$$

Ces deux termes sont manifestement égaux (il suffit d'élever et d'abaisser les mêmes indices pour passer de l'un à l'autre). En développant également le premier terme de la divergence, on a donc:

$$\partial_\alpha \theta^{\alpha\beta} = \frac{1}{\mu_0}\left[(\partial^\mu F_{\mu\lambda})F^{\lambda\beta} + F_{\mu\lambda}(\partial^\mu F^{\lambda\beta}) + \frac{1}{2}F_{\mu\lambda}(\partial^\beta F^{\mu\lambda})\right] \ . \tag{4.91}$$

Le premier terme dans le crochet peut se transformer en utilisant les équations de Maxwell $\partial^\mu F_{\mu\lambda} = \mu_0 J_\lambda$. Après une transformation triviale, on a donc:

$$\partial_\alpha \theta^{\alpha\beta} - J_\lambda F^{\lambda\beta} = \frac{1}{2\mu_0}F_{\mu\lambda}\left[\partial^\mu F^{\lambda\beta} + \partial^\mu F^{\lambda\beta} + \partial^\beta F^{\mu\lambda}\right] \ , \tag{4.92}$$

où nous avons artificiellement séparé un terme en deux. Les deux derniers termes du crochet peuvent s'écrire $-\partial^\lambda F^{\beta\mu} = \partial^\lambda F^{\mu\beta}$ en utilisant les équations de Maxwell homogènes. On a finalement:

$$\partial_\alpha \theta^{\alpha\beta} - J_\lambda F^{\lambda\beta} = \frac{1}{2\mu_0}F_{\mu\lambda}\left[\partial^\mu F^{\lambda\beta} + \partial^\lambda F^{\mu\beta}\right] \ . \tag{4.93}$$

Le second membre de cette équation est la contraction d'un tenseur antisymétrique en μ, λ avec un tenseur de rang 3, symétrique en μ, λ. Cette contraction est manifestement nulle. En écrivant le tenseur de rang 3 $K^{\mu\lambda\beta}$, on a en effet $F_{\mu\lambda}K^{\mu\lambda\beta} = F_{\lambda\mu}K^{\lambda\mu\beta}$, puisque les noms des indices muets sont indifférents. De plus, en raison des propriétés de symétrie des tenseurs, ces deux termes égaux sont opposés. Ils sont donc bien nuls. Finalement, on a

$$\partial_\alpha \theta^{\alpha\beta} = -F^{\beta\lambda}J_\lambda \ , \tag{4.94}$$

une équation 4–vectorielle. Pour en interpréter la signification physique, nous allons en écrire séparément la composante temporelle et les composantes spatiales et exprimer ces quantités en fonction des densités d'énergie, vecteur de Poynting et tenseur de Maxwell.

Composante temporelle: conservation de l'énergie

La composante temporelle $\beta = 0$ s'écrit simplement:

$$\partial_\alpha \theta^{\alpha 0} = -F^{0\lambda}J_\lambda \ . \tag{4.95}$$

Il est facile de vérifier que $F^{0\lambda}J_\lambda = \mathbf{j}\cdot\mathbf{E}/c$. En regroupant les termes, on trouve l'équation scalaire:

$$\mathbf{j}\cdot\mathbf{E} + \frac{\partial u}{\partial t} + \boldsymbol{\nabla}\cdot\boldsymbol{\Pi} = 0 \ . \tag{4.96}$$

Nous retrouvons ici l'équation de conservation de l'énergie électromagnétique que l'on établit dans les cours élémentaires à partir des équations de Maxwell. Rappelons que $\mathbf{j}\cdot\mathbf{E}$ est la densité de puissance cédée par le champ à la matière[5]. On peut écrire une équation bilan globale en intégrant l'équation précédente sur un volume $\mathcal{V}$ bordé par une surface fermée $\mathcal{S}$. On a alors:

$$\int_{\mathcal{V}} \mathbf{j}\cdot\mathbf{E}\,d\tau + \frac{dU}{dt} + \int_{\mathcal{S}} \boldsymbol{\Pi}\cdot d\mathbf{S} = 0 \ , \tag{4.97}$$

où $U = \int_{\mathcal{V}} u$ est l'énergie électromagnétique totale dans le volume $\mathcal{V}$. On trouve donc que la dérivée de l'énergie électromagnétique par rapport au temps est égale à l'opposé de la puissance totale cédée à la matière plus le flux entrant du vecteur de Poynting à travers $\mathcal{S}$. Le vecteur de Poynting décrit donc bien le transport d'énergie électromagnétique.

[5]Cette expression est, comme nous l'avons vu plus haut, correcte même si le mouvement des particules est relativiste.

Composantes spatiales: conservation de la quantité de mouvement

En remarquant que $F^{i\lambda}J_\lambda = -[\rho\mathbf{E}_i + (\mathbf{j}\times\mathbf{B})_i] = -\mathbf{f}_i$ n'est autre, au signe près, que la composante i de la densité de force de Lorentz $\mathbf{f}$, on met les composantes spatiales de notre équation de conservation sous la forme:

$$\frac{d\mathbf{g}}{dt} + \mathbf{f} + \boldsymbol{\nabla}\cdot\boldsymbol{\mathcal{T}} = 0 \ , \tag{4.98}$$

où la notation $\boldsymbol{\nabla}\cdot\boldsymbol{\mathcal{T}}$ désigne le champ de vecteurs $\partial_i\boldsymbol{\mathcal{T}}_{ij}$. Nous avons posé ici

$$\mathbf{g} = \frac{\mathbf{\Pi}}{c^2} \ . \tag{4.99}$$

Pour dégager plus précisément le sens physique de cette équation, nous l'intégrerons, comme dans le cas de l'énergie, sur un volume $\mathcal{V}$ bordé par une surface fermée $\mathcal{S}$. On obtient alors:

$$\frac{d\mathbf{P}_m}{dt} + \frac{d\mathbf{G}}{dt} + \int_{\mathcal{S}} \boldsymbol{\mathcal{T}}\cdot d\mathbf{S} = 0 \ , \tag{4.100}$$

où la notation "produit scalaire" $\boldsymbol{\mathcal{T}}\cdot d\mathbf{S}$ désigne la contraction du tenseur $\boldsymbol{\mathcal{T}}$ avec le vecteur élément de surface. Nous avons posé:

$$\frac{d\mathbf{P}_m}{dt} = \int_{\mathcal{V}} \mathbf{f} \ . \tag{4.101}$$

Ce terme représente donc la variation temporelle de la quantité de mouvement de toutes les particules matérielles contenues dans $\mathcal{V}$. Il est alors facile d'interpréter le terme $d\mathbf{G}/dt$ avec

$$\mathbf{G} = \int_{\mathcal{V}} \mathbf{g} \tag{4.102}$$

comme la variation dans le temps de la quantité de mouvement totale du rayonnement électromagnétique. Le champ $\mathbf{g}$ représente alors simplement la densité locale de quantité de mouvement du champ. Le bilan de quantité de mouvement apparaît alors de manière transparente si nous interprétons le dernier terme de l'équation (4.100) comme la quantité de mouvement sortant par unité de temps du volume $\mathcal{V}$. Cette quantité de mouvement, vectorielle, apparaît bien comme le flux à travers $\mathcal{S}$ d'une quantité tensorielle.

Nous venons bien d'établir ici le bilan de quantité de mouvement pour l'ensemble du champ et des particules chargées. Notons que, comme dans le cas du bilan d'énergie, on peut établir cette équation bilan à partir des équations de Maxwell et de la force de Lorentz. On pourra s'en convaincre aisément à titre d'exercice. Il suffit d'écrire la densité de force de Lorentz $\mathbf{f}$ en remplaçant ρ et $\mathbf{j}$ par leurs expressions en termes de $\mathbf{E}$ et $\mathbf{B}$ extraites des équations de Maxwell. Des manipulations algébriques peu agréables permettent alors de mettre le résultat sous la forme de la somme d'une dérivée partielle par rapport au temps et d'une divergence de quantité tensorielle. Une simple identification redonne alors les équations précédentes.

Notons que la quantité de mouvement du champ électromagnétique est simplement, à un facteur dimensionnel près, le vecteur de Poynting qui décrit le déplacement de l'énergie électromagnétique. Nous pouvons, qualitativement, comprendre ce résultat en termes de photons. Nous avons déjà évoqué le fait que le rayonnement se comporte tout autant comme le phénomène ondulatoire que nous traitons ici que comme un flux de particules de masse nulle, d'énergie $h\nu$, où h est la constante de Planck et ν la fréquence du rayonnement que nous supposerons pour un moment monochromatique. Le vecteur de Poynting décrit alors le flux de photons. Chaque photon se déplaçant à la vitesse c, la densité numérique de photons N devra être de l'ordre de $(1/h\nu)\Pi/c$ (nous préciserons cette discussion très qualitative au prochain paragraphe dans le cas très simple de l'onde plane). L'impulsion de chaque photon (de masse nulle) étant $h\nu/c$, la densité d'impulsion est $Nh\nu/c = \Pi/c^2$. On retrouve bien l'expression de g.

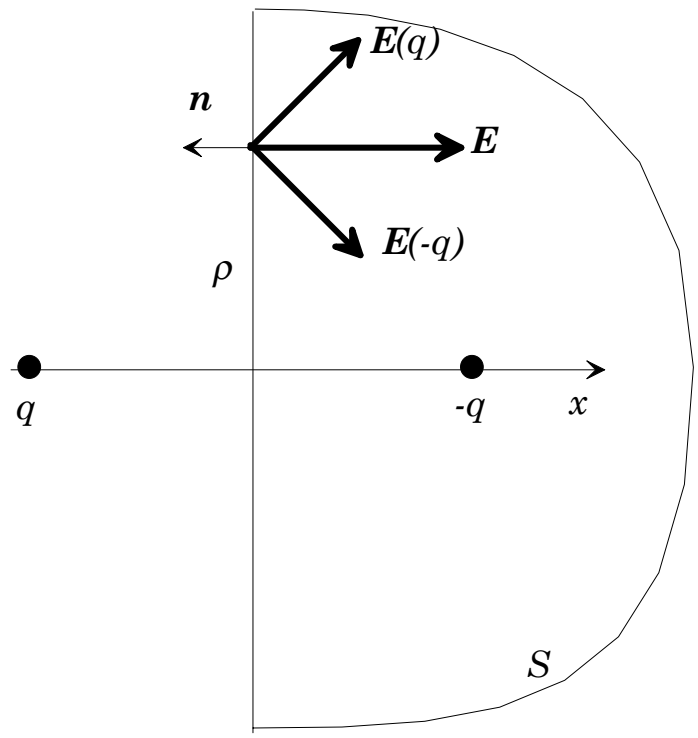

Figure 4.1: La force de Coulomb entre deux charges ponctuelles par le bilan de quantité de mouvement. Deux charges opposées sont disposées symétriquement par rapport au plan zOy. Le champ résultant est, sur ce plan, parallèle à Ox. On établit le bilan d'impulsion dans le demi-espace $x > 0$, limité par la surface S.

4.3.3 Applications

Un problème d'électrostatique

Nous allons appliquer les bilans d'énergie–quantité de mouvement à quelques situations simples. Nous allons d'abord montrer que l'équation bilan de quantité de mouvement contient simplement la force de Coulomb. Ce calcul montre que ces équations bilans peuvent présenter un intérêt en dehors des phénomènes purement propagatifs. La figure 4.1 illustre notre problème. Deux charges q et $-q$ sont situées sur l'axe Ox en $-a$ et a respectivement. Il n'échappera à personne que le calcul direct de la force de Coulomb entre ces particules ne présente aucune difficulté. Nous allons aborder ce problème par une voie un peu plus difficile. Ecrivons l'équation bilan de quantité de mouvement pour le volume $\mathcal{V}$ correspondant au demi-espace $x > 0$. Dans tout ce volume, le champ magnétique est nul. La densité de quantité de mouvement du champ est donc nulle. La variation de la quantité de mouvement de la matière, $d\mathbf{P}_m/dt$, est simplement celle de la charge $-q$ et coïncide avec la force de Coulomb $\mathbf{F}$ subie par cette charge. Notons que le système de charges en l'absence d'autres forces n'est manifestement pas en équilibre. On peut donc écrire la force de Coulomb sous la forme:

$$\mathbf{F} = -\int_{\mathcal{S}} \boldsymbol{\mathcal{T}} \cdot d\mathbf{S} \ . \tag{4.103}$$

La surface $\mathcal{S}$ bordant le volume $\mathcal{V}$ est simplement constituée du plan $x = 0$ et d'une demi sphère à l'infini. Le champ total étant à grande distance celui d'un dipôle, il décroît avec la distance R comme $1/R^3$. Le tenseur de Maxwell, proportionnel au carré du champ, décroît comme $1/R^6$. L'intégrale sur la demi sphère de l'infini est donc nulle. Sur le plan médiateur, le champ électrique est dirigé selon $\mathbf{u}_x$: $\mathbf{E} = E(\rho)\mathbf{u}_x$. Son module à une distance ρ de l'axe vaut $E(\rho) = (q/2\pi\epsilon_0)a/(a^2 + \rho^2)^{3/2}$. Il est facile alors de montrer que, $d\mathbf{S}$ étant orienté selon x, la seule composante du tenseur de Maxwell qui

joue un rôle est $\mathcal{T}_{xx} = -\epsilon_0 E(\rho)^2/2$. Le flux de $\boldsymbol{\mathcal{T}}$ sur $\mathcal{S}$ se calcule alors par une intégration triviale en coordonnées polaires. On en déduit la force de Coulomb, orientée selon $-x$, dont le module a bien sûr la valeur requise. Cet exercice facile d'électrostatique nous fait comprendre l'intérêt des équations bilans d'impulsion, même dans un cas où il n'y a pas d'effets propagatifs.

Cas de l'onde plane monochromatique

Nous allons maintenant écrire les bilans d'énergie et de quantité de mouvement dans le cas beaucoup plus important de l'onde plane. Nous allons en particulier pouvoir préciser nos interprétations en termes de photons.

Nous considérons donc une onde plane monochromatique, polarisée linéairement. Elle se propage selon Oz, le champ électrique est aligné selon Ox. La pulsation de l'onde est ω. Les champs électriques et magnétiques s'écrivent donc:

$$\mathbf{E} = E_0 \mathbf{u}_x \cos(kz - \omega t) \tag{4.104}$$

$$\mathbf{B} = \frac{E_0}{c} \mathbf{u}_y \cos(kz - \omega t) \tag{4.105}$$

avec $\omega = ck$ Le vecteur de Poynting est alors dirigé selon u_z (l'énergie se propage manifestement dans cette direction) et vaut:

$$\mathbf{\Pi} = \epsilon_0 c E_0^2 \cos^2(kz - \omega t) \mathbf{u}_z \ . \tag{4.106}$$

On ne s'intéresse souvent qu'à la moyenne temporelle de ce vecteur (les oscillations à la fréquence 2ω n'étant pas détectables, le plus souvent, dans des mesures énergétiques). On a bien sûr:

$$\overline{\mathbf{\Pi}} = \frac{\epsilon_0 E_0^2 c}{2} \mathbf{u}_z \ . \tag{4.107}$$

La valeur moyenne de la densité d'impulsion est donc:

$$\overline{\mathbf{g}} = \frac{\epsilon_0 E_0^2}{2c} \mathbf{u}_z \tag{4.108}$$

La densité d'énergie électrique est égale à la densité d'énergie magnétique. La densité d'énergie électromagnétique instantanée s'écrit donc:

$$u = \epsilon_0 E_0^2 \cos^2(kz - \omega t) \ , \tag{4.109}$$

et sa valeur moyenne vaut:

$$\overline{u} = \frac{\epsilon_0 E_0^2}{2} \ . \tag{4.110}$$

On remarque immédiatement que

$$\overline{u} = \frac{\overline{\Pi}}{c} \tag{4.111}$$

$$\overline{g} = \frac{\overline{u}}{c} \ . \tag{4.112}$$

Nous pouvons comprendre quantitativement ces relations importantes en termes de densité numérique de photons. Si nous avons N photons par unité de volume, la quantité moyenne d'énergie traversant par unité de temps une surface d'aire S perpendiculaire à l'axe de propagation sera égale d'une part à $\overline{\Pi} S$ et d'autre part à $\overline{u} c S$ (en un laps de temps dt un "volume" d'onde égal à $cdtS$ "traverse" la surface). On obtient ainsi la première relation: $\overline{u} = \overline{\Pi}/c$. La densité numérique N de photons est $N = \overline{u}/\hbar\omega$. La densité d'impulsion $N\hbar\omega/c$ est donc bien $\overline{u}/c$.

Exprimons enfin le tenseur de Maxwell. Les écritures ne présentent aucune difficulté. On trouve que seule la composante $\mathcal{T}_{zz}$ est non nulle et vaut précisément u (en valeur instantanée et donc aussi en valeur moyenne). L'onde transmet donc dans la direction z une quantité de mouvement elle aussi alignée dans la direction z. Précisons ce bilan d'impulsion en considérant le cas d'un diaphragme parfaitement absorbant d'aire S perpendiculaire à l'axe Oz. Le seul effet de ce diaphragme est d'annuler l'onde incidente immédiatement derrière lui (à plus grande distance, la diffraction par les bords du cylindre jouera un rôle et le champ électromagnétique ne sera pas exactement nul dans l'ombre géométrique du disque.

Considérons donc un volume $\mathcal{V}$ limité par une surface $\mathcal{S}$ entourant immédiatement le diaphragme. En première approximation, on pourra considérer que le champ est nul sur la "face arrière" de $\mathcal{S}$ et que le champ sur la face "avant" est celui de l'onde plane non perturbée (c'est sans doute une approximation correcte si le diaphragme est parfaitement absorbant. C'est grossièrement faux s'il est réfléchissant. Nous laissons au lecteur le soin d'examiner ce dernier cas).

Ecrivons le bilan de quantité de mouvement pour le volume $\mathcal{V}$. Comme le volume de $\mathcal{V}$, qui entoure exactement le diaphragme est négligeable, la variation de la quantité de mouvement du champ est nulle. La variation de la quantité de mouvement de la matière doit être égale à la force $\mathbf{F}$ subie par le diaphragme. Cette force est donc égale à l'opposé du flux sortant du tenseur de Maxwell. Le flux est nul partout sauf sur la surface avant du diaphragme. Il vaut alors simplement $-uS$. La force subie par le diaphragme est donc finalement

$$\mathbf{F} = uS\mathbf{u}_z \ . \tag{4.113}$$

L'onde électromagnétique exerce donc sur le diaphragme une pression p (la force est proportionnelle à la surface). Cette pression p est égale à la densité d'énergie du champ:

$$p = u \tag{4.114}$$

(cette relation tient entre valeurs instantanées, mais n'a d'intérêt qu'entre valeurs moyennes). En quelque sorte, nous écrivons là une équation d'état pour le rayonnement électromagnétique. On pourra se convaincre, là encore, que cette pression s'interprète en termes de collisions inélastiques des photons incidents avec le diaphragme.

Cette pression de radiation a de nombreuses manifestations. Notons tout d'abord qu'elle est en général assez faible. Si nous considérons, par exemple, un faisceau laser transportant 1W sur une surface de 1 mm^2, le vecteur de Poynting moyen vaut évidemment 10^6 W/mm^2, la densité d'énergie et donc la pression valent 3.10^{-3} Pa. Même si la puissance est importante, la densité d'énergie est petite, parce que la vitesse de propagation est grande. Si faible soit elle, la pression de radiation permet de faire léviter des particules suffisamment petites dans un faisceau laser. Le poids d'une particule de rayon r est en effet proportionnel à r^3 alors que la force de pression de radiation varie comme r^2. Pour r suffisamment petit, la pression de radiation l'emporte. En appliquant ce raisonnement à la pression de radiation du rayonnement solaire, on trouve que des particules suffisamment petites doivent être éjectées du système solaire. La taille limite se trouve être indépendante de la distance R au soleil. Pression de radiation et force de gravitation varient en effet toutes deux comme $1/R^2$. Avec l'ordre de grandeur de la puissance du rayonnement solaire (1.5 kW/m^2) et de l'accélération de pesanteur solaire au niveau de la terre (10^{-2} g), on trouve que des particules de rayon inférieur à 0.1 μm sont éjectées. Notons que, dans un modèle réaliste, il faudrait tenir compte aussi des collisions avec les particules chargées constituant le "vent solaire".

Partie III

Propagation, diffraction

Introduction

Nous aborderons dans ce chapitre un certain nombre de propriétés liées à la propagation des champs. Un des aspects physiquement les plus intéressants et les plus riches d'applications concerne les phénomènes de diffraction (transmission des champs électromagnétiques à travers des structures dont la taille n'est pas très grande devant la longueur d'onde). Un des objets essentiels de cette partie du cours sera donc de donner une assise rigoureuse à la théorie de la diffraction. Nous allons voir comment et au prix de quelles approximations on peut justifier les propriétés de la diffraction telles qu'elles sont enseignées dans les classes élémentaires. Nous aurons, pour cela, à utiliser des techniques très puissantes, comme celle des fonctions de Green, qui, sous d'autres formes ou d'autres noms, sont très généralement utilisées en physique.

Cette partie sera divisée en trois chapitres principaux. Dans le premier, nous établirons de façon rigoureuse la solution des équations de Maxwell en termes de potentiels retardés. Nous établirons pour ce faire la forme de la fonction de Green du champ électromagnétique, réponse des équations de Maxwell à une perturbation impulsionnelle, spatialement et temporellement. Nous intuiterons d'abord la forme de cette fonction en n'utilisant que des arguments physiques simples avant de la retrouver de façon rigoureuse. Ce chapitre fera un usage extensif de l'intégration de fonctions complexes, avec laquelle il est donc préférable d'être déjà familier.

Dans le deuxième chapitre, nous utiliserons cette solution et la fonction de Green du champ pour démontrer une formule rigoureuse (formule de Kirchhoff) relative à la propagation du champ électromagnétique. Nous utiliserons ensuite cette formule pour essayer de traiter un problème de diffraction générique, la transmission d'un champ à travers une ouverture dans un écran opaque. Nous verrons qu'on ne peut en général traiter le problème de façon rigoureuse. On devrait pour cela tenir compte des courants induits dans l'écran, ce qui est impossible. Nous ferons donc une approximation (qui est équivalente au principe de Huygens des sources secondaires) a priori très grossière mais très réaliste dans des problèmes concrets. Nous montrerons alors que le champ transmis est explicitement calculable. En faisant enfin une approximation paraxiale, nous retrouverons la diffraction dans l'approximation de Fraunhofer, telle qu'elle est enseignée généralement de façon élémentaire. Nous pourrons alors généraliser les résultats obtenus pour un écran opaque percé à des écrans semi–transparents ou même à des objets de phase.

Dans le troisième et dernier chapitre, enfin, nous traiterons brièvement et qualitativement un certain nombre d'applications pratiques de la diffraction. Nous ne ferons qu'évoquer brièvement le principe de ces techniques, sans jamais entrer dans le détail des calculs. Nous nous pencherons en particulier sur les applications de la diffraction pour le traitement optique des signaux. Nous consacrerons quelque temps à la très belle méthode de Labeyrie qui permet de s'affranchir largement des fluctuations de l'atmosphère terrestre et de rendre aux instruments astronomiques une partie de leur pouvoir de résolution théorique sans pour autant recourir à des moyens spatiaux. Nous évoquerons très brièvement le principe de l'holographie et de ses généralisations en optique non linéaire (la conjugaison de phase, par exemple). Nous montrerons enfin qualitativement, à partir de la formule de Kirchhoff, comment l'optique géométrique peut émerger du cadre de la théorie de la diffraction. Ces arguments qualitatifs seront repris sur une base beaucoup plus rigoureuse dans le quatrième appendice à cette partie.

Cette partie sera en effet prolongée de plusieurs appendices, dans lesquels on a regroupé des sujets qui touchent de près ou de loin au problème de la propagation. Il va de soi que ces appendices peuvent être omis en première lecture.

Le premier appendice, très bref, sera un rappel des propriétés de jauge du champ électromagnétique et des jauges communément utilisées. Nous insisterons en particulier sur la jauge de Coulomb qui n'est pas covariante au contraire de la jauge de Lorentz mais se trouve très utilisée dans le domaine de la physique atomique.

Dans le deuxième appendice, nous nous préoccuperons de l'écriture des champs électromagnétiques dans l'espace réciproque. Nous séparerons, dans le champ, des parties longitudinales et transverses. Nous définirons les "variables normales" associées aux champs électromagnétiques dans le vide. Nous montrerons enfin que la dynamique de ces variables normales est celle d'un simple oscillateur harmonique. Nous aurons ainsi largement préparé la voie à la quantification du champ électromagnétique qui est essentiellement une quantification d'un ensemble d'oscillateurs harmoniques correspondant aux variables normales. Cette quantification sera à peine évoquée ici mais nous en aurons donné les briques essentielles, du moins pour ce qui concerne la partie électromagnétique du problème.

Dans le troisième appendice, nous nous pencherons sur le problème de la propagation des faisceaux laser. En utilisant les résultats généraux sur la diffraction, nous construirons des faisceaux "gaussiens", solution approximative des équations de Maxwell dans l'espace libre. Ils représentent de façon assez précise, dans le cadre d'une approximation paraxiale, les faisceaux directifs et localisés des lasers. Nous établirons quelques propriétés de ces faisceaux, ainsi que quelques résultats élémentaires d'optique gaussienne.

Enfin, dans le dernier appendice, nous nous occuperons du passage de l'électromagnétisme à l'optique géométrique. Nous montrerons comment on peut retrouver rigoureusement les lois de l'optique géométrique à partir des équations de Maxwell à condition de renoncer à décrire les champs à une échelle de l'ordre de la longueur d'onde. Nous préciserons ainsi les raisonnements qualitatifs donnés à la fin du dernier chapitre.

Chapitre 1

Potentiels retardés

Nous allons, dans ce chapitre, établir rigoureusement la solution des équations de Maxwell en termes de potentiels retardés. Nous introduirons pour cela la fonction de Green du champ électromagnétique. Nous établirons l'expression de la fonction de Green dans un référentiel donné, renonçant, pour un temps, à la covariance manifeste. Nous montrerons rapidement, à la fin de ce chapitre, que la forme ainsi obtenue peut très simplement être mise sous une forme covariante, adaptée au changement de référentiel.

1.1 Fonction de Green

1.1.1 Position du problème

Notre problème est donc de résoudre les équations aux potentiels. En choisissant la jauge de Lorentz:

$$\partial_\mu \mathcal{A}^\mu = 0 = \frac{1}{c^2}\frac{\partial V}{\partial t} + \boldsymbol{\nabla}\cdot\mathbf{A} \ , \tag{1.1}$$

les équations reliant les potentiels aux sources s'écrivent simplement en termes des potentiels scalaire et vecteur:

$$\Box\,\mathbf{A} = \mu_0 \mathbf{j} \tag{1.2}$$

$$\Box\, V = \rho/\epsilon_0 \ . \tag{1.3}$$

$\Box$ est ici l'opérateur d'alembertien:

$$\Box = \partial_\mu \partial^\mu = \frac{1}{c^2}\frac{\partial^2}{\partial t^2} - \Delta \ , \tag{1.4}$$

où Δ est le laplacien.

La forme de l'équation à résoudre est donc la même pour les trois composantes du potentiel vecteur et le potentiel scalaire. Nous résoudrons donc en fait dans ce paragraphe l'équation scalaire générique:

$$\Box\,\phi(\mathbf{r},t) = \mathcal{S}(\mathbf{r},t) \ , \tag{1.5}$$

où ϕ représente une des composantes du potentiel et $\mathcal{S}$ la source associée. Notons bien sûr que cette séparation n'a de sens qu'en jauge de Lorentz. La jauge de Coulomb, par exemple, $\boldsymbol{\nabla}\cdot\mathbf{A} = 0$, est beaucoup plus ennuyeuse puisque les équations définissant $\mathbf{A}$ et V sont couplées (voir appendice 1).

1.1.2 Définition de la fonction de Green

Plutôt que de résoudre directement l'équation (1.5), cherchons une fonction de la position et du temps, $G(\mathbf{r}, t)$ telle que:

$$\Box\, G(\mathbf{r}, t) = \delta(\mathbf{r})\delta(t) \ . \tag{1.6}$$

De manière évidente, G est le potentiel rayonné par une source parfaitement localisée à l'origine de l'espace et du temps. C'est en quelque sorte la réponse percussionnelle de l'espace. Nous appellerons G la fonction de Green du potentiel.

Il est bien connu, dans la théorie de la réponse linéaire, que la réponse d'un système linéaire à une source quelconque est connue si on connaît la réponse percussionnelle. Dans le domaine du filtrage, par exemple, si on note $g(t)$ la réponse percussionnelle du filtre (la réponse quand il est sollicité par une fonction de Dirac à l'origine des temps), la réponse $s(t)$ à un signal d'entrée $e(t)$ est simplement la convolution du signal d'entrée par la réponse percussionnelle:

$$s(t) = \int_{-\infty}^{\infty} g(t-\tau)e(\tau)\, d\tau = \int_{-\infty}^{\infty} g(\tau)e(t-\tau)\, d\tau \ , \tag{1.7}$$

(l'intégrale s'étend de $-\infty$ à ∞ à condition d'admettre que la réponse percussionnelle $g(t)$ est nulle pour $t < 0$, une simple conséquence de la causalité). Cette convolution résulte de la linéarité du système et de la possibilité de développer le signal d'entrée sur les fonctions de Dirac:

$$e(t) = \int_{-\infty}^{\infty} e(\tau)\delta(t-\tau)\, d\tau \ . \tag{1.8}$$

Ces résultats sont largement utilisés, en électronique par exemple. Une détermination de la réponse percussionnelle permet de caractériser complètement un amplificateur. Notons d'ailleurs que, par une transformée de Fourier élémentaire, on trouve que la composante de Fourier à la fréquence ω du signal de sortie est proportionnelle à la composante de Fourier à la même fréquence du signal d'entrée. Le coefficient de proportionnalité, appelé susceptibilité du système à cette fréquence, n'est autre que la composante de Fourier à ω de la réponse percussionnelle. Susceptibilité et réponse percussionnelle sont donc reliées simplement par une transformation de Fourier. Nous aurons l'occasion de revenir largement sur ces problèmes dans la dernière partie de ce cours, quand nous traiterons la susceptibilité linéaire d'un milieu matériel.

Nous allons montrer explicitement que la fonction de Green, la réponse percussionnelle, nous donne aussi la solution générale de l'équation aux potentiels. Si G obéit à l'équation (1.6), on a aussi:

$$\Box_{\mathbf{r},t} G(\mathbf{r}-\mathbf{r}_1, t-t_1) = \delta(\mathbf{r}-\mathbf{r}_1)\delta(t-t_1) \ . \tag{1.9}$$

Les indices portant sur le d'alembertien précisent sur quelles variables portent les dérivations. On en déduit immédiatement:

$$\Box_{\mathbf{r},t} G(\mathbf{r}-\mathbf{r}_1, t-t_1)\mathcal{S}(\mathbf{r}_1, t_1) = \delta(\mathbf{r}-\mathbf{r}_1)\delta(t-t_1)\mathcal{S}(\mathbf{r}_1, t_1) \ , \tag{1.10}$$

les termes en $\mathcal{S}$ étant des constantes vis à vis des dérivations dans le d'alembertien. En intégrant alors les deux membres de cette équation sur toutes les valeurs d'espace et de temps de $\mathbf{r}_1$ et t_1 et en remarquant que l'intégration sur le second membre donne trivialement $\mathcal{S}(\mathbf{r}, t)$ en raison de la présence des Dirac, on obtient:

$$\mathcal{S}(\mathbf{r}, t) = \int d^3\mathbf{r}_1 dt_1 \ \Box_{\mathbf{r},t} \left[G(\mathbf{r}-\mathbf{r}_1, t-t_1)\mathcal{S}(\mathbf{r}_1, t_1)\right] \ . \tag{1.11}$$

En remarquant finalement que l'opérateur d'alembertien du membre de gauche commute avec l'intégration, on constate que le potentiel:

$$\phi(\mathbf{r}, t) = \int d^3\mathbf{r}_1 dt_1 G(\mathbf{r}-\mathbf{r}_1, t-t_1)\mathcal{S}(\mathbf{r}_1, t_1) \tag{1.12}$$

est effectivement une solution de l'équation (1.5).

Comme pour la théorie du signal élémentaire, la solution de l'équation aux potentiels est la convolution des termes sources avec la réponse percussionnelle du système. Résoudre l'équation aux potentiels est donc équivalent à trouver l'expression de la fonction de Green. Deux approches sont possibles à ce stade. Dans la première, comme nous le verrons dans le paragraphe suivant, on peut, en se fondant sur des arguments très généraux, écrire à priori la forme de la fonction de Green. Il ne reste alors qu'un coefficient arbitraire, qu'on peut identifier en essayant cette solution dans l'équation (1.6). Dans la deuxième approche, plus satisfaisante intellectuellement et à peine plus difficile, on résoudra directement cette équation en utilisant les propriétés de la transformation de Fourier.

1.1.3 Approche qualitative

Nous cherchons ici à donner, par des arguments très simples, la forme de la fonction de Green. Rappelons qu'il s'agit du potentiel rayonné par une source qui est localisée à l'origine et qui n'existe qu'à l'instant origine. En raison de l'invariance de l'espace par rotation, la fonction G doit être, comme sa source, à symétrie sphérique. Elle ne peut donc dépendre que de la distance r à l'origine.

Nous "savons bien" (ce sera l'objet du prochain paragraphe de le montrer rigoureusement) que les solutions à l'équation des potentiels décrivent des ondes se propageant à la vitesse c. La source étant nulle en tout instant sauf 0, la fonction G doit décrire une onde sphérique très localisée divergeant de l'origine à la vitesse c à partir de $t = 0$. Elle n'est donc non nulle que pour des distances telles que $t = r/c$. G est "donc" proportionnelle à $\delta(r - ct)$.

L'énergie doit se conserver. Les champs sont linéaires dans la fonction de Green. Le vecteur de Poynting est donc quadratique dans la fonction de Green. Pour que l'énergie totale transportée par cette onde soit indépendante du temps, il faut que le vecteur de Poynting soit en $1/r^2$ au niveau du "front d'onde". La fonction de Green doit donc être proportionnelle à $1/r$:

$$G(r,t) \simeq \frac{1}{r}\delta(r - ct) \ . \tag{1.13}$$

Il nous faut enfin tenir compte de la causalité. La fonction de Green doit être nulle en tous les instants précédant l'instant origine. On peut en tenir compte en ajoutant une fonction de Heaviside du temps $\theta(t)$. On a alors tenu compte de toutes les propriétés essentielles de la fonction de Green, qui doit pouvoir s'écrire:

$$G(r,t) = A\frac{1}{r}\theta(t)\delta(r - ct) \ . \tag{1.14}$$

La constante A est a priori arbitraire. Pour l'identifier, on portera cette expression dans l'équation (1.6)[1]. On trouvera:

$$A = \frac{c}{4\pi} \ . \tag{1.15}$$

Bien sûr, ce raisonnement qualitatif, bien qu'il donne le résultat exact, est insatisfaisant. Ce n'est que par substitution dans l'équation initiale qu'on peut vérifier que la solution est effectivement convenable. Nous allons consacrer le prochain paragraphe à une solution plus rigoureuse. En particulier, nous n'aurons pas à admettre la propagation des solutions à la vitesse c qui est sans doute l'hypothèse la moins justifiée a priori dans le raisonnement précédent.

[1]Nous conseillons au lecteur d'effectuer ce calcul, qui impose une manipulation soigneuse des distributions. On y comprendra mieux en particulier quel rôle joue la fonction $\theta(t)$ dans la fonction de Green. On verra qu'elle est indispensable pour obtenir un comportement convenable au voisinage de l'origine.

1.2 Solution rigoureuse

1.2.1 Fonction de Green

Pour résoudre (1.6), nous allons passer dans l'espace réciproque. Nous introduisons donc la transformée de Fourier quadridimensionnelle de G définie par:

$$\tilde{G}(\mathbf{k},\omega) = \frac{1}{(2\pi)^2}\int G(\mathbf{r},t)e^{-i(\mathbf{k}\cdot\mathbf{r}-\omega t)}\, d^3\mathbf{r}dt \ , \tag{1.16}$$

où l'intégration doit s'effectuer sur tout l'espace–temps. La relation réciproque s'écrit simplement:

$$G(\mathbf{r},t) = \frac{1}{(2\pi)^2}\int \tilde{G}(\mathbf{k},\omega)e^{i(\mathbf{k}\cdot\mathbf{r}-\omega t)}\, d^3\mathbf{k}d\omega \ . \tag{1.17}$$

L'avantage de la transformation de Fourier est de rendre triviaux les opérateurs différentiels sur l'espace ou le temps. En prenant la transformée de Fourier du premier membre de (1.6), on obtient simplement $(k^2-\omega^2/c^2)\tilde{G}$. La transformée de Fourier des fonctions de Dirac est une constante et l'équation (1.6) se met finalement dans l'espace réciproque sous la forme:

$$(k^2-\omega^2/c^2)\tilde{G}(\mathbf{k},\omega) = \frac{1}{4\pi^2} \ . \tag{1.18}$$

La solution de cette équation algébrique est élémentaire:

$$\tilde{G}(\mathbf{k},\omega) = \frac{c^2}{4\pi^2}\frac{1}{c^2k^2-\omega^2} \ . \tag{1.19}$$

Le problème est donc immédiatement résolu. Pour pouvoir utiliser cette solution, il faut toutefois repasser dans l'espace réel par une opération de transformation de Fourier inverse qui, comme nous allons le voir, est un peu plus complexe.

La fonction de Green dans l'espace réel s'écrit donc:

$$G(\mathbf{r},t) = \frac{c^2}{16\pi^4}\int \frac{e^{i(\mathbf{k}\cdot\mathbf{r}-\omega t)}}{c^2k^2-\omega^2}\, d^3\mathbf{k}d\omega \ . \tag{1.20}$$

Il est clair cette intégration n'est pas élémentaire, puisque l'intégrande présente des pôles en $\omega = \pm ck$. On pourra d'ailleurs constater tout de suite que ces pôles coïncident avec la relation de dispersion pour des ondes planes dans le vide, ce qui n'est pas tout à fait un hasard. Pour régler ce problème, nous procéderons en deux temps, d'abord à l'intégration sur les fréquences puis à l'intégration sur les vecteurs d'onde.

Nous avons donc à calculer:

$$I = \int -\frac{e^{-i\omega t}}{(\omega-ck)(\omega+ck)}\, d\omega \ , \tag{1.21}$$

l'intégration s'effectuant sur tout l'axe réel. La fonction à intégrer étant définie et analytique sur l'ensemble du plan complexe, nous allons procéder, comme il est d'usage en pareil cas, à une déformation du contour d'intégration. Au lieu d'intégrer sur l'axe réel, nous choisirons d'intégrer sur une droite parallèle, correspondant à des valeurs de ω de partie imaginaire constante égale à ϵ. Si tout se passe bien, nous pourrons prendre à la fin des calculs la limite pour $\epsilon \to 0$ et obtenir un résultat physique[2]. Nous choisirons $\epsilon > 0$. Nous discuterons dans un moment ce qu'on obtiendrait en faisant l'autre choix possible.

Pour appliquer le théorème des résidus, il nous faut fermer le contour d'intégration par un "demi–cercle à l'infini" situé soit dans le demi plan supérieur, soit dans le demi plan inférieur. Cette fermeture

[2]Nous verrons en fait que le résultat final est indépendant de ϵ, rendant trivial le passage à la limite.

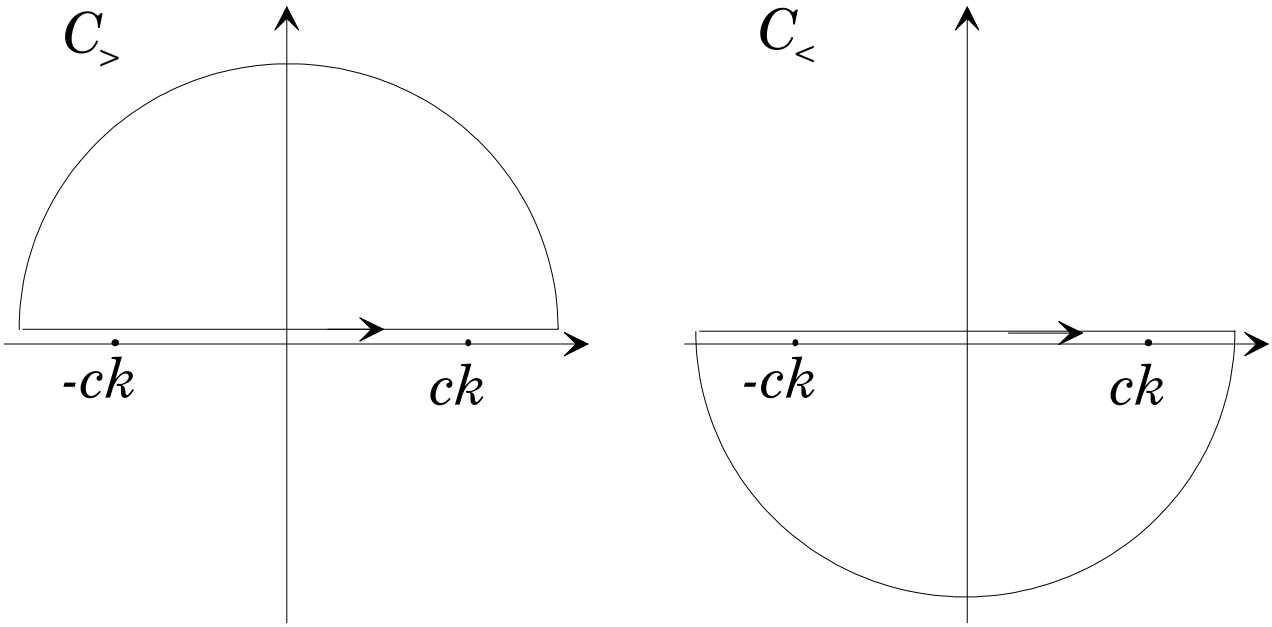

Figure 1.1: Contours d'intégration dans le plan complexe des fréquences pour le calcul de l'intégrale sur la fréquence ω intervenant dans la fonction de Green. Pour des temps négatifs, le contour de gauche convient. Pour des temps positifs, on choisira le contour de droite. Les pôles de la fonction à intégrer sont situés en $\pm ck$ sur l'axe réel.

est possible à condition que cette nouvelle partie du contour ne contribue pas à l'intégrale. Si t est positif, $e^{-i\omega t}$ tend vers 0 pour des points à distance infinie dans le demi plan inférieur, correspondant à des parties imaginaires négatives pour ω. Pour $t > 0$, nous intégrerons donc sur le contour $\mathcal{C}_<$ représenté à droite sur la figure 1.1. En revanche, pour $t < 0$, nous intégrerons sur le contour $\mathcal{C}_>$ qui se referme par un demi–cercle dans le demi plan supérieur. Les pôles de la fonction à intégrer étant tous deux sur l'axe réel, nous trouvons donc immédiatement que l'intégrale et, donc, la fonction de Green sont nulles pour les instants négatifs. G est donc proportionnelle à la fonction de Heaviside $\theta(t)$. C'est une simple expression de la causalité, la source de ce potentiel n'existant qu'à l'instant origine.

Si nous avions choisi initialement une partie imaginaire ϵ négative pour notre contour d'intégration, les choix de demi–cercles pour fermer le contour d'intégration en eussent été inchangés. En revanche, comme on s'en persuadera aisément, on aurait trouvé une fonction de Green non nulle pour $t < 0$, nulle pour $t > 0$: un potentiel qui précède la source. Une telle situation viole manifestement la causalité relativiste et ne peut convenir. Le choix d'une partie imaginaire positive pour ω n'est donc pas du tout arbitraire[3]. On pourra, à titre d'exercice, poursuivre le calcul avec $\epsilon < 0$. Sans surprise, on trouvera, au lieu des potentiels retardés en $r - ct$, la solution non physique en potentiels "avancés" en $r + ct$. La présence de ces solutions parasites est clairement due au fait que l'équation de propagation est du second ordre en t et se trouve donc invariante dans un renversement du sens du temps.

Nous n'avons donc plus à calculer I que pour $t > 0$. Les résidus de la fonction sont:

$$-\frac{e^{-ickt}}{2ck} \qquad \text{en} \qquad \omega = ck \tag{1.22}$$

$$\frac{e^{ickt}}{2ck} \qquad \text{en} \qquad \omega = -ck\ . \tag{1.23}$$

$$\tag{1.24}$$

Compte tenu du sens dans lequel est parcouru $\mathcal{C}_<$, I est égale à $-2i\pi$ fois la somme des résidus. On

[3]Nous rencontrerons le même genre d'arguments dans la partie sur l'électromagnétisme dans la matière quand nous établirons les relations de Kramers–Kronig.

trouve donc immédiatement, en regroupant les termes:

$$I = \frac{2\pi}{ck} \sin ckt \ . \tag{1.25}$$

En reportant cette expression dans celle de la fonction de Green, on obtient:

$$G(\mathbf{r}, t) = \frac{c}{8\pi^3} \int \ e^{i\mathbf{k}\cdot\mathbf{r}} \frac{\sin ckt}{k} \, d^3\mathbf{k} \ . \tag{1.26}$$

Pour procéder à cette dernière intégration spatiale, nous utiliserons les coordonnées sphériques (k, θ, ϕ) pour $\mathbf{k}$ (l'intégrande est évidemment à symétrie sphérique, ce qui reflète l'invariance par rotation). L'axe vertical de ce repère sera choisi aligné avec $\mathbf{r}$. Dans ce calcul, $\mathbf{r}$ apparaît comme une constante et ce choix est possible pour chaque position. L'intégrale sur ϕ est triviale et donne un simple facteur 2π. En remarquant que $\mathbf{k} \cdot \mathbf{r} = kr \cos\theta$, l'intégrale sur θ s'écrit simplement:

$$\int_0^\pi \sin\theta e^{ikr\cos\theta} \, d\theta = \frac{2}{kr} \sin kr \ . \tag{1.27}$$

Il ne nous reste finalement que l'intégrale sur r à calculer:

$$G(\mathbf{r}, t) = \frac{c}{2\pi^2} \frac{1}{r} \int_0^\infty \sin ckt \sin kr \, dk \ . \tag{1.28}$$

En transformant le produit de sinus en une somme de cosinus et en passant aux exponentielles complexes, on a:

$$\begin{aligned} G(\mathbf{r}, t) &= \frac{c}{8\pi^2 r} \int_0^\infty \left[e^{ik(r-ct)} + e^{-ik(r-ct)} - e^{ik(r+ct)} - e^{-ik(r+ct)} \right] dk \\ &= \frac{c}{8\pi^2 r} \int_{-\infty}^\infty \left[e^{ik(r-ct)} - e^{ik(r+ct)} \right] dk \ . \end{aligned} \tag{1.29}$$

Pour passer à la seconde ligne, on a remarqué que les exponentielles complexes de la première peuvent être regroupées deux par deux en procédant au changement de variable $k \to -k$ et en étendant donc l'intégration à tout l'axe réel. Les intégrales d'exponentielles complexes sur tout l'axe donnent simplement, à un facteur 2π près, des fonctions de Dirac et on a:

$$G(\mathbf{r}, t) = \frac{c}{4\pi} \frac{1}{r} (\delta(r - ct) - \delta(r + ct)) \ . \tag{1.30}$$

Nous ne devons pas oublier à ce stade que le calcul n'a été effectué que pour $t > 0$ et que la fonction de Green est en fait proportionnelle à $\theta(t)$. Pour des valeurs positives de t, la seconde fonction de Dirac dans l'expression précédente ne joue aucun rôle et nous pouvons finalement écrire:

$$G(\mathbf{r}, t) = \frac{c}{4\pi} \frac{1}{r} \theta(t) \delta(r - ct) \ . \tag{1.31}$$

Nous avons maintenant une forme explicite de la fonction de Green. Avant d'aller plus loin, nous allons la mettre sous une forme légèrement différente, parfois mieux adaptée à certains calculs (nous l'utiliserons en particulier pour établir la forme des potentiels retardés). Nous allons en effet montrer qu'on peut aussi écrire:

$$G(\mathbf{r}, t) = \frac{1}{4\pi} \frac{1}{r} \theta(r) \delta(t - r/c) \ . \tag{1.32}$$

Pour établir très simplement ce résultat, il suffit de remarquer que:

$$\delta(r - ct) = \delta(c(t - r/c)) = \frac{1}{c} \delta(t - r/c) \ . \tag{1.33}$$

1.2.2 Forme covariante

Nous avons établi les formes précédentes de la fonction de Green en distinguant bien espace et temps. Nous avons donc a priori perdu la covariance manifeste des équations de Maxwell. Il n'en est heureusement rien en fait et on peut, par de simples manipulations algébriques, mettre la fonction de Green sous une forme manifestement covariante.

Nous pouvons en effet écrire:

$$\theta(t)\delta(r-ct) = \theta(t)\left[\delta(r-ct) - \delta(r+ct)\right] \ . \tag{1.34}$$

Nous avons en effet déjà utilisé cette propriété (la fonction de Heaviside annule l'action de la seconde fonction de Dirac) dans l'établissement de la forme finale de la fonction de Green. La somme de fonctions de Dirac peut être réécrite:

$$\delta(r-ct) - \delta(r+ct) = 2r\delta((ct-r)(ct+r)) \ . \tag{1.35}$$

Pour nous en convaincre, nous pouvons considérer l'action de $\delta((ct-r)(ct+r))$ sur une fonction régulière $f(r,t)$. Deux points contribuent à l'intégrale sur r: $\pm ct$. Au voisinage du point $r = ct$, la fonction de Dirac, considérée comme distribution sur t, est équivalente à $\delta(2r(ct-r)) = \delta(ct-r)/2r$. Cette équivalence, ajoutée à la contribution de $r = -ct$, établit finalement cette propriété utile.

En remarquant que $c^2t^2 - r^2 = x_\mu x^\mu$, où les x^μ sont les coordonnées contravariantes du point où est calculée la fonction de Green, cette dernière peut s'écrire:

$$G(x^\mu) = \frac{c}{2\pi}\theta(x^0)\delta(x_\mu x^\mu) \ . \tag{1.36}$$

La fonction de Dirac au second membre est manifestement covariante. La fonction de Heaviside l'est aussi en dépit des apparences. Nous avons vu en effet que les transformations de Lorentz n'affectaient pas l'ordre temporel de deux événements situés dans ou sur le cône de lumière l'un de l'autre: passé, futur et causalité sont des invariants relativistes. Le signe de x^0 est donc un invariant relativiste. Nous avons donc obtenu, avec (1.36), une forme manifestement covariante de la fonction de Green.

1.2.3 Potentiels retardés

La dernière étape à franchir est élémentaire. Il ne nous reste qu'à reporter la fonction de Green dans l'expression (1.12). Nous utiliserons pour cela la deuxième forme de la fonction de Green. On a alors la solution de l'équation aux potentiels sous la forme:

$$\phi(\mathbf{r},t) = \frac{1}{4\pi}\int \theta(|\mathbf{r}-\mathbf{r}_1|)\frac{\delta(t-t_1-|\mathbf{r}-\mathbf{r}_1|/c)}{|\mathbf{r}-\mathbf{r}_1|}\mathcal{S}(\mathbf{r}_1,t_1)\,d^3\mathbf{r}_1\,dt_1 \ . \tag{1.37}$$

Nous pouvons à ce point oublier la fonction θ. Elle ne joue de rôle que pour des points sources $\mathbf{r}_1$ infiniment proches du point d'observation $\mathbf{r}$. Si la source est une répartition continue de charges, la contribution correspondante est infinitésimale. L'intégration sur le temps est alors triviale. Le seul instant t_1 qui contribue est tel que la source et le point d'observation soient sur le cône de lumière l'un de l'autre (nous reverrons plus tard ce raisonnement sous une forme un peu différente en examinant le rayonnement des charges en mouvement). On a donc:

$$\phi(\mathbf{r},t) = \frac{1}{4\pi}\int \frac{\mathcal{S}(\mathbf{r}_1,t-|\mathbf{r}-\mathbf{r}_1|/c)}{|\mathbf{r}-\mathbf{r}_1|}\,d^3\mathbf{r}_1 \ , \tag{1.38}$$

forme standard de la solution en termes de potentiels retardés. En ajoutant les facteurs dimensionnels, on retrouve en effet les formes habituelles pour les expressions des potentiels scalaire et vecteur:

$$V(\mathbf{r},t) = \frac{1}{4\pi\epsilon_0}\int \frac{\rho(\mathbf{r}_1,t-|\mathbf{r}-\mathbf{r}_1|/c)}{|\mathbf{r}-\mathbf{r}_1|}\,d^3\mathbf{r}_1 \tag{1.39}$$

$$\mathbf{A}(\mathbf{r},t) = \frac{\mu_0}{4\pi}\int \frac{\mathbf{j}(\mathbf{r}_1,t-|\mathbf{r}-\mathbf{r}_1|/c)}{|\mathbf{r}-\mathbf{r}_1|}\,d^3\mathbf{r}_1 \ . \tag{1.40}$$

L'interprétation physique de ces expressions est si transparente que nous n'y reviendrons pas.

Chapitre 2

Diffraction: approche théorique

Introduction

Un problème générique de diffraction est représenté sur la figure 2.1. Une source de rayonnement éclaire une portion d'espace à travers un écran opaque percé d'une ouverture. Nous cherchons à déterminer les champs ou les potentiels en présence de l'écran à partir de ce qu'ils sont dans l'espace libre. Bien sûr, le problème admet de nombreuses variantes, avec des écrans semi–transparents ou des écrans ne modifiant que la phase des champs, mais le principe du calcul reste toujours le même. Le volume dans lequel nous désirons déterminer le champ sera limité par une surface fermée S (la fermeture de la surface pourra éventuellement se faire par une portion de "sphère de l'infini"). Il est assez naturel que S coïncide avec l'écran là où il est présent. Nous allons en fait procéder en plusieurs étapes.

Nous établirons d'abord la formule de Kirchhoff qui permet de calculer le champ (ou plutôt, comme dans le chapitre précédent, une des composantes du potentiel) en tous points du volume $\mathcal{V}$ intérieur à S si on connaît le champ en tous points de la surface S à tous les instants. Cette formule sera établie sur des bases très générales à partir de la fonction de Green. Bien sûr, elle permettrait de résoudre directement notre problème de diffraction si on connaissait le champ sur S.

En fait, le champ de la source est clairement nul à l'infini. Il est nul aussi en tous points de l'écran, à condition de prendre S infiniment proche de l'écran du côté opposé à la source. Cela résulte de l'hypothèse que l'écran est totalement absorbant. Le problème serait donc résolu si on connaissait le champ sur la petite portion de S au voisinage immédiat du "trou" percé dans l'écran. C'est bien sûr là que réside la difficulté essentielle. Le champ dans cette région est en effet la somme du champ produit par la source et du champ rayonné par les courants induits dans l'écran. Tenir compte convenablement de ceux-ci, c'est résoudre explicitement les équations de Maxwell avec les conditions aux limites imposées par l'écran. Ce calcul est en général tout à fait impossible[1]. Il nous faut donc faire une approximation pour pouvoir continuer.

Cette approximation, équivalente au principe des ondelettes de Huygens, fera l'objet du second paragraphe. Nous assimilerons simplement le champ sur S "dans le trou" au champ libre des sources en l'absence d'écran. C'est évidemment une approximation très grossière. Nous pourrons cependant en justifier la validité dans la plupart des cas utiles et nous verrons dans la suite qu'elle rend étonnamment bien compte des phénomènes de diffraction les plus courants. Nous établirons alors une formule donnant de façon assez générale le champ diffracté dans le volume $\mathcal{V}$.

Dans le dernier paragraphe de ce chapitre, nous retrouverons la formulation élémentaire de la théorie de la diffraction en termes de transformée de Fourier de la fonction de transparence de l'écran. Nous verrons que ce régime s'établit simplement en effectuant une approximation paraxi-

[1] Un seul cas a été résolu explicitement par Sommerfeld, au début du siècle, celui de la diffraction par le bord d'un demi–plan infiniment mince. On trouvera le calcul détaillé dans le Born et Wolf.

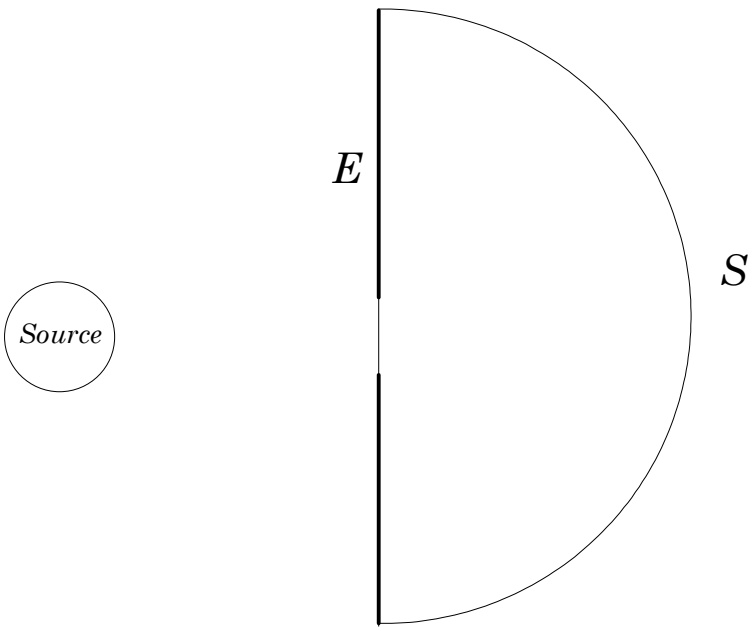

Figure 2.1: Situation générique d'un problème de diffraction. Une source éclaire une portion de l'espace, limitée par une surface S, à travers une ouverture dans un écran opaque E.

ale sur l'expression obtenue au paragraphe précédent. Nous rappellerons alors brièvement un certain nombre de résultats standard relatifs à ce régime qui nous seront utiles dans une discussion qualitative des applications de la diffraction.

2.1 Formule de Kirchhoff

La situation qui nous préoccupe ici est donc celle représentée sur la figure 2.1 *en l'absence de l'écran* E. En d'autres termes, nous voulons calculer le champ en tous points du volume $\mathcal{V}$ intérieur à la surface S à partir de la valeur du champ sur cette surface. Nous supposerons qu'il n'existe pas de sources à l'intérieur du volume $\mathcal{V}$ où nous calculons le champ. Si c'était le cas, il faudrait ajouter à notre résultat leur contribution, calculée par la formule des potentiels retardés. Il faut bien dire qu'on n'éclaire pas en général l'image de diffraction avec une lampe de poche placée derrière l'écran percé!

Nous considérerons bien sûr la forme générique de l'équation de propagation et, comme au chapitre précédent, nous noterons ϕ la composante du potentiel que nous appellerons systématiquement "champ" par abus de langage et S sa source. Nous calculerons le champ en un point générique $\mathbf{r}, t$ du volume $\mathcal{V}$. Dans toute la suite, $\mathbf{r}_1, t_1$ désignera un point de la surface S ou du volume $\mathcal{V}$ à un autre instant.

Comme il n'y a pas de sources dans $\mathcal{V}$, on a

$$\Box_1 \phi(\mathbf{r}_1, t_1) = 0 \ , \tag{2.1}$$

où $\Box_1$ désigne l'opérateur d'alembertien portant sur les coordonnées spatio–temporelles du point $\mathbf{r}_1, t_1$. On a de même, en utilisant la définition de la fonction de Green:

$$\Box_1 G(\mathbf{r} - \mathbf{r}_1, t - t_1) = \delta(\mathbf{r} - \mathbf{r}_1)\delta(t - t_1) \tag{2.2}$$

(nous dérivons ici par rapport à $\mathbf{r}_1$ et non par rapport à $\mathbf{r}$, mais le résultat est le même puisque $\Box$ ne fait intervenir que des dérivées du second ordre, insensibles au signe de l'argument). On déduit immédiatement de ces deux expressions que:

$$G \Box_1 \phi - \phi \Box_1 G = -\phi(\mathbf{r}_1, t_1)\delta(\mathbf{r} - \mathbf{r}_1)\delta(t - t_1) \ . \tag{2.3}$$

Intégrons maintenant les deux membres de cette équation sur le temps t_1 et sur les points $\mathbf{r}_1$ du volume $\mathcal{V}$. L'intégrale du second membre donne trivialement $-\phi(\mathbf{r},t)$. On a donc:

$$-\phi(\mathbf{r},t) = \int_{\mathcal{V}} d^3\mathbf{r}_1 \int_{-\infty}^{\infty} dt_1 \left[G \Box_1 \phi - \phi \Box_1 G\right] \tag{2.4}$$

Nous chercherons maintenant à transformer l'intégrale de volume en une intégrale sur la surface S. Pour cela, remplaçons maintenant le d'alembertien par son expression explicite: $\Box_1 = 1/c^2\partial^2/\partial t_1^2 - \Delta_1$. L'intégrale fait alors intervenir une somme de deux termes, l'un contenant les dérivées temporelles et l'autre les opérateurs laplaciens. Les termes temporels s'écrivent:

$$G\frac{\partial^2\phi}{\partial t_1^2} - \phi\frac{\partial^2 G}{\partial t_1^2} = \frac{\partial}{\partial t_1}\left[G\frac{\partial\phi}{\partial t_1} - \phi\frac{\partial G}{\partial t_1}\right] \tag{2.5}$$

L'intégration temporelle est donc immédiate. Elle donne un terme proportionnel à ϕ et G, évalué à $t_1 = -\infty$ ou $t_1 = \infty$. Très clairement, tout modèle raisonnable doit supposer que les champs s'annulent à des instants infiniment lointains. Ce terme tout intégré est donc manifestement nul et il ne reste que les termes en laplacien:

$$\phi(\mathbf{r},t) = \int_{\mathcal{V}} d^3\mathbf{r}_1 \int_{-\infty}^{\infty} dt_1 (G\Delta_1\phi - \phi\Delta_1 G) \ . \tag{2.6}$$

Le terme à intégrer est manifestement une divergence:

$$G\Delta_1\phi - \phi\Delta_1 G = \boldsymbol{\nabla}_1 \cdot [G\boldsymbol{\nabla}_1\phi - \phi\boldsymbol{\nabla}_1 G] \ . \tag{2.7}$$

On peut immédiatement transformer l'intégrale de volume en une intégrale sur S et écrire:

$$\phi(\mathbf{r},t) = \int_{-\infty}^{\infty} dt_1 \int_S d\mathbf{S}_1 \cdot [G\boldsymbol{\nabla}_1\phi - \phi\boldsymbol{\nabla}_1 G] \ , \tag{2.8}$$

où $d\mathbf{S}_1$ est l'élément de surface de S (avec l'orientation classique de la normale sortante). Une simple convention de notations nous permet alors d'obtenir la formule de Kirchhoff sous sa forme standard. Nous définirons la "dérivée normale" d'une fonction ϕ sur la surface S comme:

$$\frac{\partial\phi}{\partial n} = \mathbf{n} \cdot \boldsymbol{\nabla}_1\phi \ , \tag{2.9}$$

où $\mathbf{n}$ est le vecteur unitaire normal à S. Remarquant que $d\mathbf{S}_1 = dS_1\mathbf{n}$, nous avons finalement:

$$\phi(\mathbf{r},t) = \int_{-\infty}^{\infty} dt_1 \int_S \left[G\frac{\partial\phi}{\partial n} - \phi\frac{\partial G}{\partial n}\right] dS_1 \ . \tag{2.10}$$

Nous avons donc rempli nos objectifs en exprimant le champ en tous points intérieurs à $\mathcal{V}$ en fonction du champ uniquement sur S. Bien sûr, la connaissance du champ sur S à tous les instants est nécessaire pour reconstituer tout le champ dans $\mathcal{V}$. La formule de Kirchhoff ainsi établie est parfaitement générale et exacte. Notons qu'il en existe d'autres versions beaucoup plus sophistiquées, permettant par exemple de traiter les fonctions de corrélation du champ. On les trouvera décrites dans le Born et Wolf.

2.2 Principe de Huygens

Si la formule de Kirchhoff résout parfaitement la question en l'absence d'écran, elle ne nous est pas directement utile pour la diffraction. En effet, dans le cas de la figure 2.1, il est clair que le champ ϕ est nul en tous les points de S qui sont directement situés derrière l'écran. Il est nul également

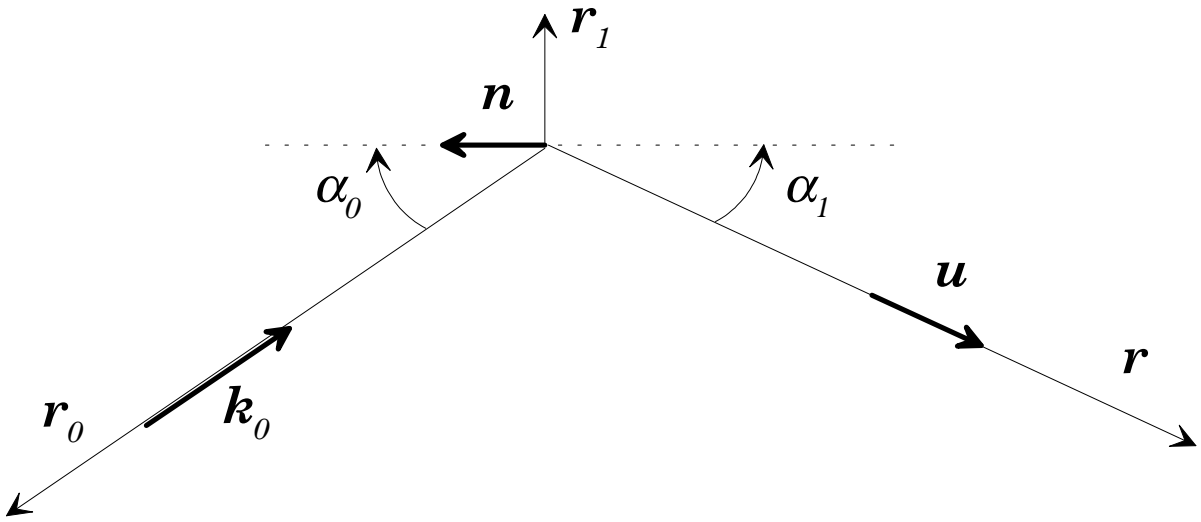

Figure 2.2: Conventions d'orientation et notations pour le calcul du champ diffracté.

sur les portions de S situées à l'infini. Il n'est non nul en fait que sur les portions de S qui sont situées immédiatement derrière le trou dans l'écran E. Le problème est que le champ à ce niveau est la somme des champs produits par les sources et des champs rayonnés par les courants induits dans l'écran. Ceux-ci, comme nous l'avons déjà vu, ne sont pas en général calculables. Il nous faut donc recourir à une approximation pour poursuivre.

Cette approximation est plutôt brutale. Nous allons supposer que *les champs en tous points de la surface S derrière le trou sont identiques aux champs rayonnés par les sources en l'absence de l'écran.* Nous négligeons ainsi les courants induits dans l'écran qui sont précisément responsables de son opacité! Nous pouvons cependant essayer d'imaginer dans quels types de conditions cette approximation pourrait être valable.

Elle est d'abord loin d'être valable si les dimensions du trou sont de l'ordre de la longueur d'onde ou plus petites. Il est clair que, dans une telle structure, les champs sont considérablement modifiés par rapport à ce qu'ils sont dans l'espace libre[2]. Il est clair aussi que cette approximation tombe si l'épaisseur de l'écran est beaucoup plus grande que les dimensions transverses du trou. Là aussi, les champs à la sortie n'ont rien à voir avec les champs directement rayonnés par les sources. Si cette approximation grossière a quelques chances de refléter la réalité, c'est pour des ouvertures ayant une extension grande devant la longueur d'onde et percées dans des écrans minces. Heureusement, c'est une situation qu'on rencontre souvent en optique. En fait, cette approximation est bien meilleure en pratique que ne le laisse supposer son aspect.

Nous allons pouvoir alors expliciter le calcul des dérivées de la fonction de Green et du champ et terminer le calcul de la formule de Kirchhoff. Pour simplifier encore un peu, nous allons supposer (ce qui est en général le cas en pratique) que les sources sont monochromatiques et situées loin de l'écran (à l'échelle de la taille du trou). En d'autres termes, nous allons assimiler le champ des sources à une onde plane de fréquence ω au niveau du trou. Nous considérerons des points $\mathbf{r}_1$ sur la portion de la surface S au "voisinage" du trou. La figure 2.2 précise un certain nombre de conventions d'orientation.

Nous choisirons une origine O située dans cette portion de S. La normale $\mathbf{n}$ à S est orientée en direction de la source. Si le trou est assez petit, nous pourrons assimiler la portion utile de S à un

[2]Notons que, dans ce cas, le calcul des champs transmis est possible. En effet, les phénomènes de propagation jouant peu pour des dimensions plus petites que λ, on peut raisonner comme si les champs étaient statiques. On calcule donc le champ électrique, en électrostatique, à l'orifice du trou. On lui ajoute alors une dépendance sinusoïdale en temps et on peut calculer ensuite le champ rayonné, par exemple en appliquant la formule de Kirchhoff à une surface bien choisie. On trouvera des discussions de problèmes similaires dans le Jackson.

plan. $\mathbf{n}$ est donc constant. Les sources sont situées à grande distance en $\mathbf{r}_0$ dans une direction faisant un angle α_0 avec $\mathbf{n}$. Le vecteur d'onde de l'onde plane incidente sera noté $\mathbf{k}_0$. Le point d'observation $\mathbf{r}$ est situé dans une direction faisant un angle α_1 avec *l'opposé* de $\mathbf{n}$. Nous noterons $\mathbf{u}$ le vecteur unitaire de la direction d'observation. Le point courant de l'intégration est simplement repéré par sa position $\mathbf{r}_1$.

Nous avons donc essentiellement à calculer $\partial\phi/\partial n$ et $\partial G/\partial n$. Le champ des sources au point $\mathbf{r}_1$ s'écrit:

$$\phi(\mathbf{r}_1, t_1) = e^{i\mathbf{k}_0\cdot\mathbf{r}_1} e^{-i\omega t_1} \ . \tag{2.11}$$

On a donc

$$\boldsymbol{\nabla}_1\phi = i\mathbf{k}_0\phi \tag{2.12}$$

et

$$\frac{\partial\phi}{\partial n} = -ik\cos\alpha_0\phi(\mathbf{r}_1, t_1) \ , \tag{2.13}$$

où k est le module du vecteur d'onde incident $\mathbf{k}_0$. Dans ce calcul, on prendra garde aux différentes orientations.

Ecrivons maintenant $\partial G/\partial n$. Le gradient de $G(\mathbf{r}-\mathbf{r}_1, t-t_1)$ peut s'écrire, en remarquant que la fonction de Green est à symétrie sphérique par rapport au point source ou au point d'observation:

$$\boldsymbol{\nabla}_1 G = -\mathbf{u}\frac{\partial G}{\partial|\mathbf{r}_1-\mathbf{r}|} \ . \tag{2.14}$$

Nous utiliserons pour G la forme (1.32). $\mathbf{r}$ étant dans ce cas toujours très différent de $\mathbf{r}_1$, la distribution θ ne joue aucun rôle et nous pouvons l'oublier dans le calcul de la dérivée. Nous devons donc dériver $\delta(t-t_1-|\mathbf{r}-\mathbf{r}_1|/c)/4\pi|\mathbf{r}-\mathbf{r}_1|$ ce qui donne immédiatement:

$$\boldsymbol{\nabla}_1 G = \mathbf{u}\left[\frac{\delta'(t-t_1-|\mathbf{r}-\mathbf{r}_1|/c)}{4\pi c|\mathbf{r}-\mathbf{r}_1|} + \frac{\delta(t-t_1-|\mathbf{r}-\mathbf{r}_1|/c)}{4\pi|\mathbf{r}-\mathbf{r}_1|^2}\right] \ , \tag{2.15}$$

où δ' désigne la dérivée de la distribution de Dirac par rapport à son argument. Le gradient de G fait donc intervenir deux termes. L'un est en $1/|\mathbf{r}-\mathbf{r}_1|$, l'autre en $1/|\mathbf{r}-\mathbf{r}_1|^2$. Quand l'observation a lieu à grande distance, le second terme est négligeable par rapport au premier. On peut bien sûr ne pas être très convaincu par un raisonnement d'ordres de grandeur appliqué à des distributions. Notons qu'on peut rendre le raisonnement plus rigoureux en faisant agir les deux distributions présentes dans le gradient de G sur une fonction régulière du temps et de l'espace oscillant à la fréquence ω, f. Le deuxième terme donne une contribution en $f/|\mathbf{r}-\mathbf{r}_1|^2$ et le premier une contribution en $f(\omega/c)/|\mathbf{r}-\mathbf{r}_1|$. Comme ω/c n'est autre que la longueur d'onde λ du rayonnement, le premier terme est plus grand que le second par un facteur de l'ordre de $|\mathbf{r}-\mathbf{r}_1|/\lambda$. Nous pouvons donc négliger le deuxième terme dès que la distance d'observation est notablement plus grande que λ. L'écriture finale de $\partial G/\partial n$ ne pose aucun problème et on a:

$$\frac{\partial G}{\partial n} = -\frac{1}{4\pi c}\cos\alpha_1\frac{\delta'(t-t_1-|\mathbf{r}-\mathbf{r}_1|/c)}{|\mathbf{r}-\mathbf{r}_1|} \ . \tag{2.16}$$

Nous reportons alors l'expression des dérivées normales dans la formule de Kirchhoff (2.10). Nous commencerons par estimer le terme le plus complexe, en $\phi\partial G/\partial n$. Il s'écrit:

$$\begin{aligned} A &= -\int_{-\infty}^{\infty} dt_1 \int dS_1\, \phi\frac{\partial G}{\partial n} \\ &= \int dS_1\, \frac{1}{4\pi c}\cos\alpha_1\frac{1}{|\mathbf{r}-\mathbf{r}_1|}\int_{-\infty}^{\infty} dt_1\, \phi(\mathbf{r}_1, t_1)\delta'(t-t_1-|\mathbf{r}-\mathbf{r}_1|/c) \ , \end{aligned} \tag{2.17}$$

en procédant d'abord à l'intégrale sur t_1 et en en ôtant tous les termes indépendants de t_1. Dans cette expression et toutes les suivantes, l'intégrale de surface s'étend uniquement à la portion de S située

immédiatement derrière l'orifice. On pourrait étendre l'intégration à toute la surface en multipliant ϕ par une fonction de transparence, nulle partout sauf dans ce domaine. Pour ne pas alourdir inutilement les notations, nous ne procéderons à cette transformation que dans le paragraphe suivant. L'intégrale sur le temps peut se calculer par parties. Le terme tout intégré est manifestement nul en raison de la présence de la distribution de Dirac. L'intégrale sur t_1 s'écrit donc finalement

$$\int_{-\infty}^{\infty} dt_1 \frac{\partial \phi}{\partial t_1} \delta(t - t_1 - |\mathbf{r} - \mathbf{r}_1|/c) = \left.\frac{\partial \phi}{\partial t_1}\right)_{t_1 = t - |\mathbf{r} - \mathbf{r}_1|/c} . \tag{2.18}$$

ϕ étant une fonction oscillante à la fréquence ω, sa dérivée temporelle à l'instant retardé peut tout aussi bien s'écrire:

$$\left.\frac{\partial \phi}{\partial t_1}\right)_{t_1 = t - |\mathbf{r} - \mathbf{r}_1|/c} = -i\omega \phi(\mathbf{r}_1, t) e^{ik|\mathbf{r} - \mathbf{r}_1|} . \tag{2.19}$$

Remarquons que nous faisons, par un artifice trivial, intervenir ici la valeur du champ estimée au point $\mathbf{r}_1$, non pas au temps retardé t_1, mais bien au temps d'observation t. L'avantage immédiat de cette approche est que le temps d'observation, au contraire du temps retardé, est invariable quand $\mathbf{r}_1$ varie (nous n'avons toujours pas fait l'intégration spatiale). On a donc finalement:

$$A = -\frac{i\omega}{4\pi c} \int dS_1 \, \cos \alpha_1 \phi(\mathbf{r}_1, t) \frac{e^{ik|\mathbf{r} - \mathbf{r}_1|}}{|\mathbf{r} - \mathbf{r}_1|} \tag{2.20}$$

Nous calculons maintenant l'autre terme dans la formule de Kirchhoff:

$$\begin{aligned} B &= \int_{-\infty}^{\infty} dt_1 \int dS_1 \, G \frac{\partial \phi}{\partial n} \\ &= \int dS_1 \frac{-ik \cos \alpha_0}{4\pi |\mathbf{r} - \mathbf{r}_1|} \int_{-\infty}^{\infty} dt_1 \, \phi(\mathbf{r}_1, t_1) \delta(t - t_1 - |\mathbf{r} - \mathbf{r}_1|/c) . \end{aligned} \tag{2.21}$$

L'intégrale sur le temps est triviale et donne:

$$\phi(\mathbf{r}_1, t - |\mathbf{r} - \mathbf{r}_1|/c) = \phi(\mathbf{r}_1, t) e^{ik|\mathbf{r} - \mathbf{r}_1|} . \tag{2.22}$$

On a donc finalement:

$$B = -\frac{ik}{4\pi} \int dS_1 \, \cos \alpha_0 \phi(\mathbf{r}_1, t) \frac{e^{ik|\mathbf{r} - \mathbf{r}_1|}}{|\mathbf{r} - \mathbf{r}_1|} \tag{2.23}$$

En regroupant les termes A et B et en remarquant que $k = 2\pi/\lambda$ et $\omega/c = k$, on obtient finalement:

$$\phi(\mathbf{r}, t) = \frac{1}{i\lambda} \int dS_1 \, \chi \phi(\mathbf{r}_1, t) \frac{e^{ik|\mathbf{r} - \mathbf{r}_1|}}{|\mathbf{r} - \mathbf{r}_1|} , \tag{2.24}$$

avec

$$\chi = \frac{1}{2}(\cos \alpha_0 + \cos \alpha_1) . \tag{2.25}$$

L'interprétation physique de cette expression est transparente. Le champ au point d'observation s'écrit comme la somme d'ondes sphériques (le terme en e^{ikr}/r) rayonnées par les différents points du trou. Chacune de ces ondes a une amplitude proportionnelle à l'amplitude de l'onde incidente (terme en ϕ), multipliée par un terme constant en $1/i\lambda$ et par un facteur purement géométrique décrivant l'inclinaison des "rayons lumineux" par rapport à la normale, χ. Nous retrouvons ici sur une base "rigoureuse" le principe des ondelettes de Huygens. Les premières approches théoriques de la diffraction postulaient en effet que chaque élément de surface de l'ouverture diffractante était une "source secondaire" d'une "ondelette" sphérique d'amplitude proportionnelle à celle de l'onde incidente. On retrouve ainsi l'essentiel de la formule (2.24). Il n'y manque que le préfacteur qui n'a aucune importance si on ne s'intéresse finalement qu'à l'intensité diffractée en valeur relative et le

facteur χ. Celui-ci est très voisin de l'unité si l'écran est à peu près normal à la direction de l'onde incidente et de l'onde diffractée, ce qui est en général le cas dans une expérience de ce type. Notons que cette dépendance en cosinus des angles n'est pas sans rappeler la loi de Lambert d'émissivité des surfaces.

Le calcul pratique de l'intégrale est en général assez difficile, comme on pourra s'en convaincre aisément. Un des succès les plus marquants de l'optique ondulatoire, qui s'est longtemps opposée à une conception corpusculaire, a été la démonstration théorique et expérimentale, par Poisson, de la présence d'un maximum d'intensité sur l'axe, à distance finie, pour la diffraction par un obstacle circulaire. Le développement des calculs de diffraction au siècle dernier a suscité l'introduction de nombreuses fonctions spéciales, telles la célèbre fonction d'Airy. Des méthodes graphiques ont aussi été développés pour sommer les amplitudes diffractées (méthode des spirales de Cornu, par exemple). Avec les calculateurs numériques, ces méthodes sont tombées en désuétude. Elles ont cependant un contenu physique qui est loin d'être inintéressant (on pensera en particulier à la méthode des zones de Fresnel). On pourra consulter, pour en savoir plus, le Bruhat–Kastler d'optique, très complet sur ce sujet.

2.3 Diffraction de Fraunhofer

2.3.1 Approximation paraxiale

Nous restreindrons un peu la généralité du paragraphe précédent pour obtenir des expressions plus simples. Nous décrirons en fait une expérience typique de diffraction telle qu'elle est réalisée en optique. En général, l'écran est plan et la source l'éclaire normalement. On a donc $\cos\alpha_0 = 1$ et ϕ est une constante dans le domaine d'intégration que nous noterons ϕ_1. Nous ferons aussi l'hypothèse que l'observation s'effectue à très grande distance. Nous ne garderons donc dans l'expression finale que les termes au premier ordre non trivial en $1/|\mathbf{r}-\mathbf{r}_1|$. Finalement, nous supposerons que les angles de diffraction sont petits et donc que la direction d'observation est, elle aussi, pratiquement normale au plan de l'écran. On aura donc $\cos\alpha_1 = 1$ et donc $\chi = 1$. Notons que cette approximation est tout à fait consistante avec l'hypothèse d'illumination normale et le principe de Huygens qui ne tient que pour des ouvertures notablement plus grande que λ, pour lesquelles les angles de diffraction restent modérés.

En reportant ces différentes hypothèses dans (2.24), on met le champ diffracté sous la forme:

$$\phi(\mathbf{r},t) = \frac{\phi_1}{i\lambda}\int dS_1\,\frac{e^{ik|\mathbf{r}-\mathbf{r}_1|}}{|\mathbf{r}-\mathbf{r}_1|}\ . \tag{2.26}$$

L'intégrale, là encore, ne porte que sur la portion de S située immédiatement derrière l'ouverture. Comme l'observation se fait à grande distance, le terme $1/|\mathbf{r}-\mathbf{r}_1|$ varie très peu quand le point $\mathbf{r}_1$ parcourt le domaine d'intégration. Il est donc tout à fait légitime de remplacer, à l'ordre le plus bas, ce terme par $1/r$, distance de l'origine (située dans le domaine d'intégration par hypothèse) au point d'observation. Le terme de phase, en revanche, doit être traité plus soigneusement. Il peut varier notablement sur le domaine d'intégration, puisqu'une variation de l'ordre de λ de la distance suffit. Nous allons donc développer ce terme:

$$(\mathbf{r}-\mathbf{r}_1)^2 \simeq r^2(1-2\frac{\mathbf{r}\cdot\mathbf{r}_1}{r^2}) \tag{2.27}$$

et donc:

$$|\mathbf{r}-\mathbf{r}_1| \simeq r - \frac{\mathbf{r}\cdot\mathbf{r}_1}{r}\ . \tag{2.28}$$

L'exponentielle complexe peut donc s'écrire:

$$e^{ik|\mathbf{r}-\mathbf{r}_1|} = e^{ikr}e^{-i\mathbf{k}\cdot\mathbf{r}_1} \tag{2.29}$$

en posant

$$\mathbf{k} = k\frac{\mathbf{r}}{r} = k\mathbf{u} \ . \tag{2.30}$$

Nous allons également étendre formellement l'intégration à tout le plan de l'écran. Nous repérerons les positions dans l'espace par un trièdre $Oxyz$. L'origine O est la même. Les axes Ox et Oy sont dans le plan de l'écran, l'axe Oz pointe dans la direction d'observation. Pour étendre l'intégration, nous introduisons donc la *fonction de transparence* de l'écran $T(x, y)$. Elle est égale à un en tous points de l'orifice, à zéro partout ailleurs. Avec ces notations, en utilisant les résultats précédents, le champ diffracté à la distance r dans la direction $\mathbf{k}$ définie simplement par les composantes transverses k_x et k_y du vecteur $\mathbf{k}$ s'écrit:

$$\phi(k_x, k_y, r) = \frac{\phi_1}{i\lambda}\frac{e^{ikr}}{r}\int dx_1 dy_1\, T(x_1, y_1) e^{-i(k_x x_1 + k_y y_1)} \ . \tag{2.31}$$

Nous retrouvons ici le résultat central des cours élémentaires sur la diffraction: l'amplitude diffractée est proportionnelle à la transformée de Fourier bidimensionnelle de la fonction de transparence $\tilde{T}(k_x, k_y)$. Le terme nouveau dans notre approche est le préfacteur. Avant d'explorer rapidement les propriétés essentielles de la figure de diffraction, mentionnons que le calcul précédent peut être généralisé sans difficultés à une incidence arbitraire. L'onde incidente a alors une phase dépendant linéairement de la position dans l'orifice. Le résultat de ce facteur de phase additionnel est une simple translation de la transformée de Fourier. Essentiellement, la figure de diffraction est la même qu'en incidence normale mais elle se trouve centrée autour de la direction incidente (un résultat bien intuitif si on imagine une ouverture de très grande taille). Nous conseillons au lecteur d'établir les expressions correspondantes à titre d'exercice.

2.3.2 Différentes expressions de la figure de diffraction

Avant de nous pencher sur les propriétés générales de la diffraction, nous allons mettre (2.31) sous des formes équivalentes, correspondant à différents modes d'observation de la figure de diffraction.

Notons tout d'abord qu'on peut facilement exprimer ϕ en fonction des angles α et β que font les projections dans les plans xOz et yOz du vecteur $\mathbf{k}$ avec l'axe Oz:

$$\begin{aligned} \alpha &= k_x/k \\ \beta &= k_y/k \ , \end{aligned}$$

ces angles étant tous petits dans l'approximation paraxiale. On a alors:

$$\phi(\alpha, \beta, r) = \frac{\phi_1}{i\lambda}\frac{e^{ikr}}{r}\int dx_1 dy_1\, T(x_1, y_1) e^{-ik(\alpha x_1 + \beta y_1)} \ . \tag{2.32}$$

Telle que nous l'avons calculée, la figure de diffraction n'est observable qu'à l'infini (c'est à dire à une distance de l'écran grande par rapport à toutes les dimensions du problème). En pratique, on peut, par exemple, recueillir simplement la figure de diffraction sur un écran plan, perpendiculaire à Oz, situé à une distance D de l'origine. C'est la manière dont on procède quand la figure de diffraction est suffisamment lumineuse (par exemple quand on éclaire l'objet diffractant par un faisceau laser). En repérant les points de cet écran par leurs coordonnées x et y, on constate immédiatement que $k_x = kx/D$ et $k_y = ky/D$. On peut alors mettre le champ reçu au point x, y sous la forme:

$$\phi(x, y, D) = \frac{\phi_1}{i\lambda}\frac{e^{ikD}}{D}\int dx_1 dy_1\, T(x_1, y_1) e^{-i(k/D)(x x_1 + y y_1)} \ . \tag{2.33}$$

Si la figure de diffraction n'est pas très lumineuse, il peut être avantageux de l'observer à distance finie dans le plan focal d'une lentille convergente d'axe optique Oz. En repérant là encore le plan focal

par les coordonnées x et y, on montre facilement que la direction définie par α et β est conjuguée par la lentille du point $x = f\alpha$, $y = f\beta$ où f est la distance focale de la lentille. En omettant les préfacteurs (qui dépendraient dans ce cas de l'ouverture numérique de la lentille), on met alors la figure de diffraction sous la forme:

$$\phi(x, y) \propto \int dx_1 dy_1 \, T(x_1, y_1) e^{-i(k/f)(xx_1+yy_1)} \; . \tag{2.34}$$

Pour toutes ces écritures différentes, la propriété essentielle demeure: la figure de diffraction est la transformée de Fourier à deux dimensions de la fonction de transparence.

2.3.3 Généralisation à une transparence arbitraire

Telle que nous l'avons définie, la fonction de transparence vaut 0 ou 1 et correspond au cas d'un écran qui est soit totalement absorbant soit totalement transparent. En réalité, un objet diffractant peut modifier continûment l'amplitude de l'onde incidente et altérer sa phase. L'onde incidente étant plane et normale au plan de l'écran, l'amplitude en amont de l'écran s'écrit toujours ϕ_1. La quantité qui intervient dans tous nos raisonnements pour déterminer le champ diffracté est la valeur de l'amplitude sur la surface S qui est située en aval de l'écran. Si l'écran est suffisamment mince, son effet sur l'amplitude en un point est a priori une multiplication de l'amplitude incidente en ce point par un facteur réel de réduction (ou d'amplification) d'intensité et par une phase, c'est à dire par un facteur complexe. La condition pour que l'on puisse traiter aussi simplement la transmission par un objet de phase est que celui-ci soit assez mince[3]. Une lentille mince, par exemple, dans un plan situé immédiatement derrière elle, ne modifie que la phase de l'onde mais pas la répartition initiale d'amplitude. Il suffit pour cela, en termes de rayons lumineux, que la déviation d'un rayon dans la traversée de l'objet soit négligeable devant les dimensions transverses caractéristiques du problème. Une lentille mince, un verre inhomogène obéiraient à ces conditions. En revanche, une sphère diélectrique pleine ou un verre fortement dépoli (et donc constitué de micro-lentilles) ne les respecteraient pas. La fonction de transparence ainsi généralisée est donc une fonction à valeurs complexes. Toutes les expressions de la figure de diffraction obtenues dans les paragraphes précédents restent valables.

2.3.4 Propriétés générales de la figure de diffraction

Nous allons dégager dans ce paragraphe quelques propriétés générales de la figure de diffraction qui ne sont en fait qu'une réécriture des propriétés standard de la transformée de Fourier.

Etendue spatiale et angulaire

Il est bien connu que l'étendue de la transformée de Fourier est inversement proportionnelle à l'étendue de la fonction. Plus la fonction de transparence aura des variations rapides, plus elle sera localisée, plus la figure de diffraction sera ouverte. Plus précisément, si Δx_1 est l'extension caractéristique de la fonction de transparence dans la direction x_1, l'étendue de la figure de diffraction sera, en termes des différentes variables utilisées pour la décrire:

- $\Delta k_x = 1/\Delta x_1$ en termes des composantes du vecteur d'onde;
- $\Delta\alpha = 1/(k\Delta x_1)$ en termes des coordonnées angulaires;
- $\Delta x = D/(k\Delta x_1)$ pour une projection à l'infini;
- $\Delta x = f/(k\Delta x_1)$ pour l'observation dans le plan focal d'une lentille.

[3] Ce genre d'approximations est assez largement utilisé dans de nombreux domaines de la physique. On le rencontre en particulier en mécanique quantique pour décrire l'effet d'un potentiel localisé sur une onde plane incidente. Dans ce contexte, le régime qui nous intéresse porte le nom de "régime de Raman–Nath".

Diffraction et translation

Considérons deux écrans caractérisés par deux fonctions de transmission T_a et T_b. Nous supposerons que le second écran se déduit du premier par une simple translation spatiale:

$$T_b(x_1, y_1) = T_a(x_1 - x_0, y_1 - y_0) \ . \tag{2.35}$$

En notant ϕ_a et ϕ_b les deux figures de diffraction et en omettant les préfacteurs pour alléger les notations, on a:

$$\begin{aligned} \phi_b(k_x, k_y) &= \int T_b(x_1, y_1) e^{-i(k_x x_1 + k_y y_1)} \\ &= \int T_a(x_1 - x_0, y_1 - y_0) e^{-i(k_x x_1 + k_y y_1)} \\ &= e^{-i(k_x x_0 + k_y y_0)} \int T_a(x_1, y_1) e^{-i(k_x x_1 + k_y y_1)} \ . \end{aligned} \tag{2.36}$$

Les deux figures de diffraction sont donc identiques, à un facteur de phase près. A titre d'exercice, on pourra écrire ce facteur de phase pour les autres expressions de la figure de diffraction.

Théorème des écrans complémentaires

Considérons encore deux écrans, de fonctions de transmission T_a et T_b. Nous dirons que ces écrans sont *complémentaires* si:

$$T_b + T_a = 1 \ . \tag{2.37}$$

Pour des écrans absorbants percés, la complémentarité exprime simplement que tout point opaque sur l'un est transparent sur l'autre et réciproquement. Un trou circulaire est ainsi complémentaire d'un obstacle circulaire de même diamètre. Les figures de diffraction ϕ_a et ϕ_b sont simplement les transformées de Fourier de T_a et T_b. La linéarité de la transformée de Fourier et la relation de complémentarité nous permettent donc d'écrire, en omettant une fois de plus les préfacteurs:

$$\phi_a + \phi_b \propto \int e^{-i(k_x x_1 + k_y y_1)} \simeq \delta(k_x)\delta(k_y) \ . \tag{2.38}$$

On a donc:

$$\phi_a = -\phi_b \tag{2.39}$$

partout sauf dans la direction de l'axe Oz. On en déduit le théorème des écrans complémentaires ou théorème de Babinet: *les figures de diffraction de deux écrans complémentaires sont identiques sauf sur l'axe.*

La différence entre les figures de diffraction sur l'axe a une origine très claire. Dans le cas d'un diaphragme circulaire, par exemple, il y a beaucoup moins (infiniment moins) de lumière transmise dans la direction initiale de l'onde plane éclairante que pour un obstacle circulaire complémentaire. Nous verrons dans le prochain chapitre qu'on peut utiliser un diaphragme circulaire pour filtrer une image optique. On peut par exemple purifier géométriquement un faisceau laser par ce moyen. Il est évident, bien que le filtrage soit fondé sur les propriétés de la figure de diffraction, que le faisceau transmis serait très différent si on utilisait un écran complémentaire, puisqu'il n'y aurait plus de lumière!

2.3.5 Quelques exemples

Nous rappellerons très brièvement dans ce paragraphe l'expression de quelques figures de diffraction très standard, auxquelles les cours élémentaires ont du familiariser le lecteur. Si ce n'était pas le cas, nous engageons le lecteur à rétablir ces expressions.

Fente

Nous considérons d'abord une fente rectangulaire dont les côtés sont alignés avec les axes, de largeur a selon Ox, b selon Oy. Nous exprimerons la figure de diffraction ϕ_f en termes des angles α et β et nous omettrons une partie des préfacteurs. On a:

$$\phi_f(\alpha,\beta) \propto ab\frac{sin(k\alpha a/2)}{k\alpha a/2}\frac{sin(k\beta b/2)}{k\beta b/2} \ . \tag{2.40}$$

On se reportera aux manuels d'optique pour y voir cette figure de diffraction.

Fentes d'Young

Nous considérons deux fentes rectangulaires identiques à la précédente séparées par une distance ℓ le long de Ox. C'est la configuration de l'expérience des interférences d'Young, la plus simple des expériences d'interférométrie optique. La figure de diffraction ϕ_y s'obtient immédiatement en utilisant les propriétés de translation et la linéarité de la transformation. On a:

$$\phi_y(\alpha,\beta) = \phi_f(1+e^{-ik\alpha\ell}) \ , \tag{2.41}$$

où ϕ_f est la figure de diffraction d'une des fentes. On peut, à partir de là, exprimer l'intensité diffractée:

$$I(\alpha,\beta) = 4I_f\cos^2\frac{k\alpha\ell}{2} \ , \tag{2.42}$$

où I_f est l'intensité diffractée par une fente unique. La séparation entre les fentes étant bien sûr plus grande que leur largeur, la modulation due à la superposition des figures de diffraction des deux fentes est beaucoup plus rapide que l'extension de I_f. On retrouve donc des interférences modulant la figure de diffraction d'une fente unique. L'interfrange est, en termes de l'angle α, de $2\pi/k\ell = \lambda/\ell$, un résultat qu'on peut retrouver de manière élémentaire en considérant la différence de marche entre les deux rayons lumineux provenant des deux fentes.

Réseau

Le réseau généralise à N fentes disposées régulièrement les résultats précédents. On obtient, par les mêmes arguments:

$$\begin{aligned}\phi_r &\propto \phi_f\left(1+e^{-ik\alpha\ell}+e^{-i2k\alpha\ell}+\cdots+e^{-iNk\alpha\ell}\right)\\ &= \phi_f\frac{1-e^{-iNk\alpha\ell}}{1-e^{-ik\alpha\ell}}\\ &\propto \phi_f\frac{\sin kN\alpha\ell/2}{\sin k\alpha\ell/2} \ .\end{aligned} \tag{2.43}$$

Notons que nous avons éliminé un facteur de phase global sans importance physique en cours de calcul. ϕ_f est bien sûr la figure de diffraction d'une fente unique. L'allure de ϕ_r est assez complexe. Elle présente essentiellement des maxima principaux pour

$$\alpha = \frac{p\lambda}{\ell} \ , \tag{2.44}$$

où p est un entier. Ces maxima principaux sont appelés "ordres du réseau". Leur largeur est de l'ordre de $\lambda/4N\ell$. Ils sont donc beaucoup plus fins que la figure de diffraction de la fente ou que l'interfrange des fentes d'Young. Cette finesse est à l'origine du pouvoir de résolution élevé, proportionnel au nombre de traits, des réseaux utilisés en spectroscopie.

L'amplitude des maxima principaux est proportionnelle à N. L'intensité est donc quadratique dans le nombre de fentes, ce qui reflète l'addition cohérente des amplitudes issues des différentes fentes. Par "cohérente", dans ce cas, on entend que ces amplitudes sont toutes en phase (plus généralement, on réserve ce terme aux cas où les amplitudes intervenant dans une sommation ont des relations de phase).

Réseau aléatoire

Posons–nous maintenant le problème d'un réseau de N fentes réparties aléatoirement et non plus régulièrement dans le plan. Nous noterons x_i, y_i la position du centre de chaque fente. Nous ne nous préoccuperons pas d'éventuels problèmes de recouvrement, supposant que les fentes sont suffisamment petites et de densité suffisamment faible. Cette situation modélise assez bien, par exemple, la diffraction par un brouillard (sur le pont de Tolbiac) de fines gouttelettes. En appliquant les mêmes arguments que précédemment, nous trouverons:

$$\phi_b(\alpha, \beta) = \phi_f \sum_i e^{-ik(\alpha x_i + \beta y_i)} \tag{2.45}$$

ϕ_f étant encore une fois la figure de diffraction d'un objet unique. Nous pouvons clarifier un peu cette expression en calculant le carré de son module, c'est à dire l'intensité:

$$I(\alpha, \beta) = I_f \left| \sum_i e^{-ik(\alpha x_i + \beta y_i)} \right|^2 . \tag{2.46}$$

En développant le module carré dans cette expression, il apparaîtra deux types de termes. Nous aurons d'abord des termes carrés, au nombre de N, égaux tous à 1. Les doubles produits se regroupent bien sûr deux par deux pour ne donner qu'une somme de termes réels. Le nombre de ces termes est de N^2 et leur module "moyen" est de 1. Leur signe en revanche dépend des relations entre les phases. Si les ouvertures sont disposées au hasard sur l'écran, la valeur de cette somme est donc le résultat d'une marche aléatoire, ou "marche du marin ivre" le long de l'axe, avec N^2 pas de longueur moyenne 1. Il est bien connu que la distance ainsi parcourue est, en moyenne sur un grand nombre de réalisations, proportionnelle à la racine carrée du nombre de pas. La contribution des termes double produit à la somme est donc elle aussi de l'ordre de N. Au total, l'intensité diffractée est proportionnelle à N:

$$I \simeq N I_f . \tag{2.47}$$

La figure de diffraction de cet ensemble aléatoire de fentes est simplement proportionnelle, en termes d'intensité, à la figure de diffraction d'une des fentes. On a là un résultat complètement différent de celui du réseau, où l'interférence entre les différentes figures de diffraction les modifiait complètement. Nous avons en effet ici une addition *incohérente* des amplitudes diffractées par les différentes fentes, en raison du caractère aléatoire des relations de phase entre ces amplitudes. Une manifestation de cette incohérence est le fait que l'intensité résultante est proportionnelle à N. Dans le cas d'une addition incohérente, ce sont les intensités qui s'ajoutent, alors que ce sont les amplitudes dans le cas d'une diffusion cohérente, résultant alors dans une intensité proportionnelle à N^2.

Pour en revenir à une illustration physique de ce phénomène, on aura avantage à observer, à grande distance, un réverbère à travers un brouillard de densité moyenne. On observera, autour de la tache centrale, des anneaux faiblement colorés. Ils correspondent effectivement à la figure de diffraction en lumière blanche d'une gouttelette "moyenne". A partir de cette observation, on pourra s'amuser à estimer le diamètre moyen des gouttelettes.

Chapitre 3

Applications de la diffraction

Nous allons brièvement passer en revue dans ce chapitre un certain nombre d'applications de la diffraction. Nous verrons en fait qu'elle peut jouer un rôle extrêmement néfaste dans des expériences d'optique, en limitant par exemple le pouvoir de résolution de nos instruments d'optique. Elle peut, dans le même temps, constituer un outil irremplaçable de traitement du signal optique.

Nous commencerons par parler brièvement du pouvoir de résolution des instruments d'optique. C'est un sujet si vaste que nous ne ferons que l'effleurer. Nous consacrerons ensuite un paragraphe un peu plus étoffé aux techniques de traitement du signal optique. Nous décrirons en particulier la très belle méthode de Labeyrie qui permet, avec des instruments optiques basés au sol, de s'affranchir dans une large mesure de la perte de pouvoir de résolution due aux turbulences atmosphériques. Le paragraphe suivant sera consacré à une description qualitative de l'holographie et de quelques unes de ses applications. Nous dirons aussi un mot de la conjugaison de phase optique, observée en optique non linéaire, qui peut être vue comme une holographie en temps réel. Dans le dernier paragraphe de ce chapitre, nous verrons comment les notions de diffraction permettent de justifier qualitativement l'optique géométrique et la notion de rayon lumineux, ainsi que le principe de Fermat de chemin optique extrémal.

Dans tout ce chapitre, nous resterons très qualitatifs. Nous n'écrirons pratiquement jamais d'équations. Nous nous contenterons de discuter du contenu physique des effets rencontrés. En fait, aller plus loin peut entraîner à des calculs très complexes, tout à fait hors de propos ici. Par exemple, le calcul explicite du pouvoir de résolution d'un microscope optique est extrêmement complexe. On doit, en particulier, faire intervenir les propriétés de cohérence spatiale et temporelle de la lumière qui éclaire l'objet. Les lecteurs intéressés pourront trouver dans le Born et Wolf de longues discussions de ces différents problèmes.

3.1 Pouvoir de résolution des instruments d'optique

La taille finie des lentilles utilisées dans les instruments d'optique est une source de diffraction qui peut sévèrement limiter leur pouvoir de résolution. Le problème général étant complexe, nous ne discuterons qu'un cas particulier mais de grande importance pratique, celui de la lunette astronomique. Celle-ci est représentée sur la figure 3.1. Elle est constituée d'une première lentille, l'objectif, de grand diamètre D et de longue focale f. Cette lentille forme une image des objets situés à l'infini dans son plan focal. L'axe optique de la lunette étant z, nous repérerons les positions dans ce plan focal par des coordonnées x et y. Une étoile située à une distance angulaire α de l'axe (dans le plan zx pour fixer les idées) a son image en $x = -f\alpha$ dans le plan focal. Une seconde lentille de beaucoup plus courte focale f', l'oculaire, permet d'examiner "à la loupe" cette image. Le plan focal objet de l'oculaire est confondu avec le plan focal image de l'objectif. La même étoile donne donc, à la sortie de l'oculaire, un faisceau de lumière parallèle, c'est à dire une image à l'infini. L'angle sous lequel est vue cette

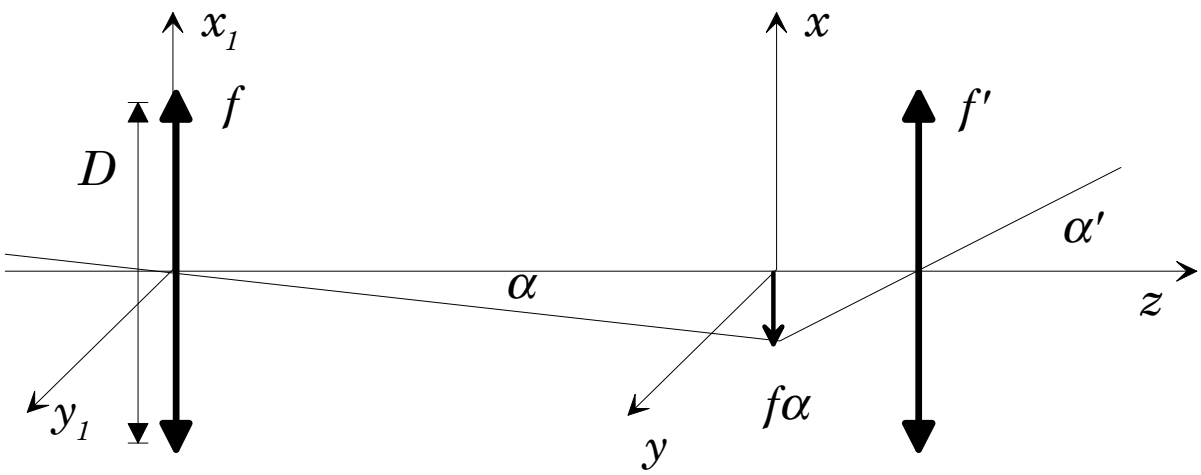

Figure 3.1: Schéma de principe d'une lunette astronomique ou d'un télescope. z est son axe optique. Le plan de la lentille d'entrée, de focale f et de diamètre D est repéré par x_1, y_1. Le plan focal de cette lentille est repéré par x, y.

image, α', est manifestement égal à $-\alpha f/f'$. Le grandissement angulaire de ce dispositif afocal est alors simplement $-f/f'$, beaucoup plus grand que 1 en module par construction.

Nous avons ainsi décrit le fonctionnement de la lunette astronomique qui n'est plus guère usitée en raison de la difficulté de produire des lentilles de très grand diamètre de qualité et de transparence suffisantes. En revanche, les télescopes les plus modernes fonctionnent exactement sur le même principe, les lentilles étant remplacées par des miroirs sphériques ou paraboliques. Notre discussion s'appliquera donc aussi bien à ces dispositifs, ainsi qu'aux jumelles, où on rajoute simplement un dispositif redresseur d'image pour obtenir une image droite.

La lumière provenant d'une étoile lointaine est, à une excellente approximation, une onde plane (nous négligerons pour le moment les problèmes de phase dûs à la traversée de l'atmosphère). L'amplitude dans le plan x_1, y_1 situé juste avant la lentille est donc une constante, si nous supposons pour un moment que l'étoile est sur l'axe optique. Nous pouvons alors voir la correspondance entre les amplitudes dans les plans x_1, y_1 et x, y comme une expérience de diffraction typique dans le régime de Fraunhofer. L'objet diffractant est l'ouverture circulaire de la lentille d'entrée (le diaphragme d'ouverture). L'objectif donne clairement une image à distance finie de la figure de diffraction. Nous pouvons donc intuiter la forme de l'image reçue dans le plan focal. Au lieu d'une image strictement ponctuelle comme en optique géométrique (nous négligerons aussi les imperfections optiques de la lentille), on obtient une petite tache dont le diamètre est de l'ordre de $f\lambda/D$, où λ est la longueur d'onde incidente. La forme détaillée de cette tache, appelée tache d'Airy, s'exprime simplement avec la fonction de Bessel d'ordre 0. On en trouvera l'expression dans le Born et Wolf. Notons seulement qu'une tache centrale, de rayon $1.22 f\lambda/D$ est entourée d'une série d'anneaux, d'intensité rapidement décroissante.

On peut estimer le pouvoir de résolution de la lunette en évaluant la distance angulaire minimale entre deux étoiles résolues dans le plan focal (on peut se convaincre aisément que l'oculaire ne limite en rien ce pouvoir de résolution si sa qualité optique est convenable). Il faut que les deux taches d'Airy soient séparées par au moins leur rayon. Il faut donc que $f\alpha > 1.22 f\lambda/D$, soit encore $\alpha > \lambda/D$. Ce pouvoir de résolution est tout à fait remarquable pour les gros instruments. Si nous prenons une longueur d'onde incidente de l'ordre du micron et un diamètre d'entrée de l'ordre du mètre, la résolution théorique serait du micro–radian, c'est à dire de l'ordre du centième de seconde d'arc.

En fait, comme nous le verrons dans un prochain paragraphe, l'effet de la turbulence atmosphérique limite le pouvoir de résolution des instruments basés au sol, même dans de très bonnes conditions atmosphériques, au dixième de seconde d'arc environ. Le pouvoir de résolution d'un grand télescope

n'est donc pas supérieur à celui d'un télescope d'amateur, avec une ouverture d'une dizaine de centimètres. L'intérêt de construire de grandes machines est autre: il est de recueillir le plus de lumière possible pour pouvoir observer des objets plus faibles et donc sans doute plus lointains. Notons aussi que certaines techniques permettent de restituer à un grand instrument une partie de son pouvoir de résolution. La première et la plus astucieuse est la méthode de Labeyrie que nous discuterons dans le prochain paragraphe. L'autre, beaucoup plus récente, est l'optique adaptative. Comme l'atmosphère déforme le front d'onde de la lumière stellaire, l'idée est de compenser cette déformation en temps réel par une déformation opposée du système optique. Il faut pour cela disposer d'une référence. Ce peut être par exemple une étoile très brillante située à proximité de l'objet à imager. On interpose dans le trajet optique un miroir déformable actionné par des cales piézo-électriques. En asservissant la position de ces cales pour minimiser la taille de l'image de l'étoile de référence, on compense les fluctuations atmosphériques et on peut obtenir un pouvoir de résolution très amélioré, comparable à celui d'un télescope spatial. Les calculateurs modernes rendent possible cette correction en temps réel, à un coût infiniment moindre qu'une expérience embarquée.

Si nous appliquons maintenant notre raisonnement à l'œil, nous pouvons estimer la limite à son pouvoir de résolution due à la diffraction par la pupille. Le diamètre de celle-ci est de l'ordre du millimètre, la longueur d'onde étant encore de l'ordre du micron. Le pouvoir de résolution angulaire devrait donc être du milli–radian. C'est effectivement, comme on pourra s'en convaincre aisément en examinant à distance les graduations d'une règle, l'ordre de grandeur du pouvoir de résolution d'un œil sain. Le plus remarquable n'est sans doute pas que nos ordres de grandeur soient corrects mais que le cristallin, lentille biologique, très adaptable, atteigne pratiquement la limite de résolution imposée par la diffraction.

Nous pouvons aussi appliquer cette discussion au pouvoir de résolution d'un appareil photographique. On l'exprime en général en "paires de traits par millimètres" dans le plan focal. En prenant le cas d'un objectif standard ($f = 50$ mm), ouvert à $f/4$, le diamètre D est de l'ordre du centimètre. Avec une longueur d'onde du micron, la limite de résolution angulaire est de 10^{-4} radians. La distance correspondante dans le plan focal est de 5 μm, correspondant à 100 paire de lignes par millimètre. On a effectivement obtenu la résolution moyenne d'un objectif de bonne qualité qui est donc bien limité seulement par la diffraction.

3.2 Traitement optique du signal

Ainsi que nous l'avons discuté dans le paragraphe précédent, la diffraction dans un instrument d'optique correspond à une opération de transformation de Fourier entre l'amplitude dans le plan de l'objectif et l'amplitude dans le plan focal. Une simple lentille est donc un calculateur analogique qui calcule, très bien, les transformées de Fourier bidimensionnelles. On peut donc l'utiliser pour réaliser des fonctions non triviales de traitement du signal.

3.2.1 Filtrage spatial

Considérons un des dispositifs optiques les plus simples, le télescope afocal représenté sur la figure 3.2. Il est identique à la lunette astronomique, si ce n'est que le grandissement vaut -1 puisque les deux lentilles, que nous continuerons à appeler objectif et oculaire pour simplifier, sont de même focale. Nous supposerons qu'un objet de transparence non uniforme est placé immédiatement avant la lentille d'entrée et nous nous intéresserons à la répartition d'amplitude sur la lentille de sortie. Ces deux répartitions d'amplitudes sont, si on néglige les imperfections dues à la diffraction par la taille finie des lentilles, exactement identiques à un grandissement total de -1 près. Si nous regardons maintenant le même dispositif, non du point de vue de l'optique géométrique, mais de celui de la diffraction, nous pourrons le trouver plus subtil qu'il n'y paraît.

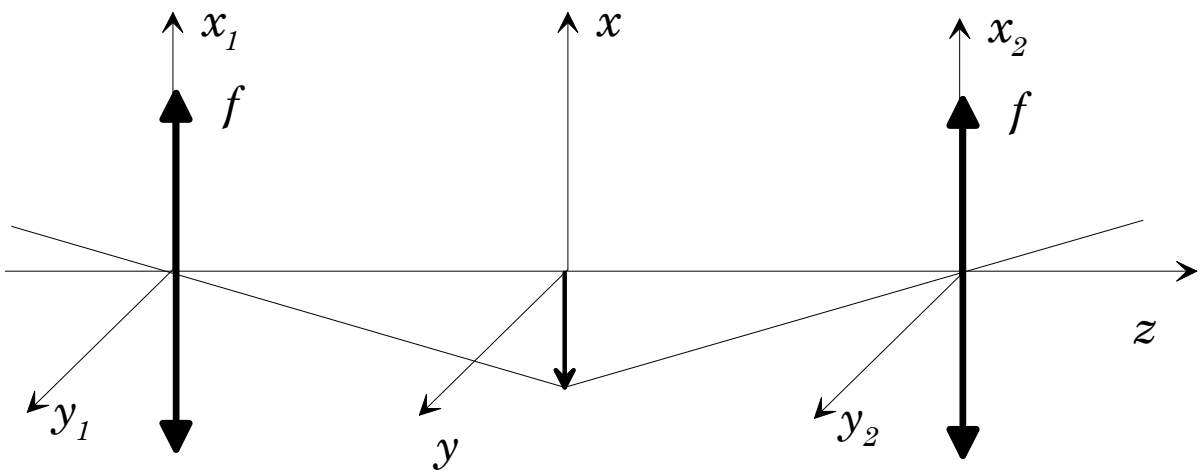

Figure 3.2: Téléscope afocal de grandissement unité. Les plans d'entrée, focal image et de sortie seront repérés respectivement par les coordonnées transverses x_1, y_1; x, y et x_2, y_2. L'axe optique est encore l'axe z.

La répartition d'amplitude dans le plan focal x, y est en effet la transformée de Fourier bidimensionnelle de la répartition d'amplitude dans le plan d'entrée x_1, y_1. L'objectif forme effectivement une image dans ce plan de la figure de diffraction de Fraunhofer. La répartition d'amplitude dans le plan de sortie est donc également la transformée de Fourier de la répartition d'amplitude dans le plan focal. En définitive, la répartition dans le plan x_2, y_2 est la double transformée de Fourier de la répartition dans le plan x_1, y_1. Elle lui est donc identique (à un changement de signe près pour l'orientation des axes en raison du grandissement négatif).

L'intérêt du dispositif est que nous disposons, dans le plan focal, de la transformée de Fourier de la répartition incidente. Nous pouvons donc, en insérant à ce niveau un diaphragme d'amplitude ou de phase, modifier cette transformée de Fourier et donc la figure de sortie. L'application la plus simple est le filtrage spatial. Supposons que le faisceau incident soit un faisceau laser gaussien (voir appendice 3). La répartition d'intensité devrait être une fonction lentement variable de la position. En raison d'effets d'interférences ou de défauts optiques, cette répartition est souvent imparfaite. Elle présente des structures à petite échelle (speckles ou tavelures) qui peuvent s'avérer très fâcheuses. Pour les éliminer, on peut insérer dans le plan focal un diaphragme de diamètre bien ajusté[1]. La répartition d'amplitude gaussienne doit en effet donner une tache de petit diamètre (essentiellement le diamètre correspondant à la limite de diffraction sur un diamètre d'entrée égal à celui du faisceau). Les tavelures et autres structures à petite échelle correspondent à des fréquences spatiales élevées et se retrouvent, dans le plan focal, à relativement grande distance de l'axe. Ces composantes de fréquence (spatiales) sont interceptées par le diaphragme. La répartition de sortie ne contient plus que les basses fréquences caractéristiques du faisceau gaussien qui se trouve donc "lissé" et débarrassé du bruit. Cette technique de "filtrage spatial" est extrêmement utile dans des expériences d'optique laser.

On peut aussi utiliser le filtrage spatial pour des problèmes d'optique plus classique. On peut ainsi, par exemple, "détramer" des photographies. Une photographie très agrandie ou transmise par des moyens numériques rudimentaires présente un tramage. L'information utile, l'image, est multipliée par une fonction maximale sur les points de la trame, nulle à mi–chemin. En un mot, l'image présente un aspect "pixellisé". Pour raisonner plus simplement, nous prendrons une image à une dimension. Imaginons qu'on ait tiré une épreuve de la photographie tramée sur film transparent. La figure utile peut alors s'exprimer comme une fonction de transparence $F(x_1)$. Elle est multipliée par le "tramage",

[1]Les résultats de l'appendice 3 permettront au lecteur intéressé de calculer le waist du faisceau dans le plan focal. Le rayon du diaphragme est typiquement choisi égal à deux fois le waist. On pourra s'amuser à estimer la perte d'intensité résultant du filtrage.

que nous prendrons comme une fonction sinusoïdale de x_1 variant entre 0 et 1, $\sin^2(x_1/x_0)$. Plaçons notre épreuve dans le plan d'entrée de notre télescope afocal et éclairons–la par une onde plane. Dans le plan focal, nous obtenons la transformée de Fourier de la fonction de transparence de l'épreuve $G(x_1) = F(x_1)\sin^2(x_1/x_0)$.

Cette transformée de Fourier est le produit de convolution de la transformée de Fourier de F et de la transformée de Fourier de la sinusoïde. Celle-ci est constitué de trois pics de Dirac. L'un, sur l'axe, correspond au fait que cette fonction a une valeur moyenne non nulle. Les deux autres correspondent à la fréquence spatiale de cette fonction (apparaissant une fois pour les fréquences positives et une fois pour les fréquences négatives, puisque la fonction d'origine est réelle). Le produit de convolution est alors une opération triviale. On obtient, dans le plan focal, trois répliques de la transformée de Fourier de F, centrées sur les trois Diracs de la transformée de la fonction de tramage. Si la fréquence spatiale du tramage est très grande par rapport aux fréquences spatiales présentes dans F, ces trois répliques ne se recouvrent pas. On peut dire aussi que l'extension spatiale d'un "pixel" doit être très petite par rapport à l'extension spatiale des détails de l'image F pour que les trois images dans le plan focal soient distinctes.

Si nous plaçons dans le plan focal un diaphragme qui ne laisse passer que la réplique centrale, nous calculerons, avec la seconde lentille, la transformé de Fourier de la transformée de F, légèrement tronquée dans l'espace des fréquences. La répartition d'amplitude en sortie ne contiendra que l'image photographique utile, débarrassée de tout tramage. Bien sûr, cela s'effectue au prix d'une perte globale d'intensité et d'une perte de résolution. Les plus petits détails visibles dans l'image de sortie sont plus grands que la taille caractéristique du tramage initial. Cette technique permet de s'affranchir d'une information parasite mais elle ne permet pas de recréer l'information perdue. Elle peut être généralisée pour effectuer d'autres traitements, plus complexes. On peut en particulier se débarrasser ainsi de certains flous photographiques (bougé, mauvaise mise au point...). Ces techniques, développées en particulier par Françon, sont maintenant un peu tombées en désuétude avec le traitement numérique du signal. Un calculateur est infiniment moins efficace qu'un afocal pour calculer une transformée de Fourier bidimensionnelle mais il est beaucoup plus versatile, évite les difficultés des plaques photographiques et permet de réaliser des calculs autres que la transformée de Fourier (les transformées en ondelettes sont de plus en plus utilisées pour les traitement de signaux complexes).

Nous avons décrit deux expériences où le filtrage diminue les fréquences spatiales présentes dans l'image. On peut l'utiliser aussi pour doubler, par exemple, une fréquence spatiale. L'expérience est un grand classique des montages d'agrégation. On dispose dans le plan d'entrée une grille de pas a (nous raisonnerons à une dimension). On observe alors dans le plan focal la transformée de Fourier de cette grille. Elle fait intervenir la fréquence spatiale fondamentale de la grille, $1/a$, et tous ses harmoniques (seulement les harmoniques impairs pour une fonction carrée de rapport cyclique 0.5). On observe donc, de part et d'autre de l'axe, une série de taches équidistantes. Insérons maintenant, dans le plan focal, une grille de pas adapté qui ne laisse passer que les harmoniques pairs de la fréquence fondamentale. L'opération de transformée de Fourier effectuée par la lentille de sortie restituera alors une fonction modulée spatialement à une fréquence *double* de la fréquence initiale! En un mot, si le pas de la grille d'entrée est a, on observera en sortie une modulation (une grille) de pas $a/2$. Voici une expérience délicate mais spectaculaire qui illustre bien les possibilités du traitement optique du signal.

3.2.2 Méthode de Labeyrie

La méthode de Labeyrie utilise astucieusement les propriétés de la transformation de Fourier pour s'affranchir, dans les instruments astronomiques basés au sol, de l'influence de la turbulence atmosphérique. Comme prototype d'instrument, nous considérerons la lunette astronomique représentée sur la figure 3.1. Si l'atmosphère n'existait pas, la lumière provenant d'un étoile lointaine située sur l'axe serait une onde plane d'amplitude constante dans le plan x_1, y_1 donnant, dans le plan focal, une

tache d'Airy dont le diamètre n'est limité que par la diffraction.

En fait, l'atmosphère perturbe les surfaces d'ondes. Son effet, si elle n'est ni trop épaisse ni trop agitée (une bonne nuit d'observation astronomique dans un très bon site), est de multiplier l'onde incidente par un facteur de phase qui peut s'écrire dans le plan d'entrée $\exp iw(x_1, y_1, t)$. La fonction w est variable dans le temps. Dans un très bon site, elle reste stable pendant quelques dizaines de millisecondes. Elle varie de 2π sur une échelle spatiale Δx_1 de l'ordre de 10 cm. La phase de l'onde étant complètement brouillée à cette échelle, on comprend bien pourquoi les plus grands télescopes n'ont pas un pouvoir de résolution meilleur que celui d'un télescope de 10 cm d'ouverture. Notons que cette vision de l'effet de l'atmosphère correspond seulement aux meilleures conditions d'observation astronomique dans un site calme et de haute altitude. Au niveau du sol, dans un lieu habituel, l'effet est beaucoup plus violent. Des effets de focalisation affectent phase et amplitude à une beaucoup plus petite échelle. Ce sont ces variations d'amplitude qui sont responsables du scintillement apparent des étoiles.

La tache observée dans le plan focal est donc la transformée de Fourier de la fonction w, tronquée au diamètre D de l'ouverture de l'objectif. C'est donc le produit de convolution d'une tache d'Airy, d'extension $f\lambda/D$, par la transformée de w. Celle-ci est très complexe, avec une taille totale de l'ordre de $f\lambda/\Delta x_1$ et des détails à toutes les échelles jusqu'à celle de la fonction d'Airy. La tache focale est donc large $(f\lambda/\Delta x_1)$ avec des détails dont la plus petite échelle est $f\lambda/D$ et change complètement au moins 100 fois par seconde. Une observation lente ne révèle donc qu'une vaste tache d'intensité uniforme: tout le pouvoir de résolution du télescope est perdu.

En revanche, si la source est suffisamment intense pour qu'on puisse enregistrer la tache de diffraction instantanée, elle révèle des speckles ou tavelures dont l'extension correspond encore au pouvoir de résolution théorique de l'instrument. On peut donc espérer, pour ces sources intenses, récupérer l'information. Pour illustrer la méthode, supposons que nous ayons à séparer deux étoiles. L'une est sur l'axe, l'autre est dans une direction inclinée, dans le plan x_1, z, d'un petit angle θ par rapport à l'axe optique. Si l'atmosphère était absente, l'amplitude dans le plan d'entrée serait une constante ϕ_1 pour la première étoile et vaudrait $\phi_2 \exp ik\theta x_1$ pour la seconde ($k = 2\pi/\lambda$). En fait, les deux ondes planes incidentes sont déformées par l'atmosphère. Si deux fonction w différentes agissaient sur les deux ondes, la situation serait sans espoir.

Examinons donc les conditions dans lesquelles les deux ondes planes issues des deux étoiles sont multipliées par le même facteur de phase. Il faut que, dans la traversée de l'atmosphère dense, les deux ondes aient traversé les mêmes fluctuations d'indice. Il faut donc que l'écart maximum entre les trajets soit de l'ordre de la distance sur laquelle w varie notablement, une dizaine de centimètres en pratique. Cet écart est $L\theta$, où L mesure l'épaisseur de l'atmosphère. Notre calcul s'appliquera donc si $L\theta < 10$ cm. En prenant $L = 10$ km, on trouve $\theta < 10^{-5}$ rd. La méthode de Labeyrie ne s'applique qu'à des étoiles séparées par une distance angulaire plus petite que 10 μrd. Mais, en raison d'une coïncidence numérique remarquable, cette distance angulaire correspond précisément au pouvoir de résolution d'un télescope limité par les fluctuations atmosphériques, $\lambda/\Delta x_1$. Dès que les taches de speckle de deux étoiles commencent à être confondues, on peut commencer à appliquer la méthode de Labeyrie. On peut donc écrire les ondes incidentes sur la lentille d'entrée sous la forme:

$$\phi_1 e^{iw(x_1,y_1,t)}$$
$$\phi_2 e^{iw(x_1,y_1,t)} e^{ik\theta x_1}$$

Les ondes provenant des deux étoiles étant manifestement incohérentes, sans aucune relation de phase, les termes d'interférences sont nuls et les éclairements produits par les deux sources dans le plan focal de l'objectif s'ajoutent simplement: $I = I_1 + I_2$ avec:

$$I_1(x,y) = \left| \int \phi_1 e^{iw(x_1,y_1,t)} e^{-i(k/f)(xx_1+yy_1)} \right|^2 \tag{3.1}$$

$$I_2(x,y) = \left| \int \phi_1 e^{iw(x_1,y_1,t)} e^{ik\theta x_1} e^{-i(k/f)(xx_1+yy_1)} \right|^2 \tag{3.2}$$

$$= \quad I_1(x - f\theta, y) \ , \tag{3.3}$$

le facteur de phase additionnel dans l'amplitude de la seconde étoile produisant une simple translation de la figure de diffraction. Notons que la translation correspond précisément à la distance entre les deux images dans le plan focal en optique géométrique. La tache obtenue dans le plan focal est donc la superposition de deux taches de speckle identiques, translatées de $f\theta$ l'une par rapport à l'autre. En corrélant la position des irrégularités de ces deux images, on pourra remonter à la distance angulaire entre les deux étoiles.

En fait, on peut très simplement extraire l'information utile par une simple expérience de diffraction. Si l'image enregistrée dans le plan focal de l'objectif est photographiée et l'épreuve développée, on obtiendra un objet dont la transmission est proportionnelle à l'intensité incidente I (pour un tirage en négatif, ce qui est en général le cas, l'opacité est proportionnelle à l'intensité incidente. Toutefois, grâce au théorème des écrans complémentaires, la figure de diffraction d'un tirage positif et d'un négatif sont identiques). Eclairons alors cet objet par une onde plane. L'amplitude diffractée dans la direction définie par les angles α et β sera donnée par:

$$g(\alpha, \beta) = \int I(x, y) e^{-ik(\alpha x + \beta y)} \, dxdy \ . \tag{3.4}$$

Elle est donc la somme des transformées de Fourier de I_1 et I_2. Si $g_1(\alpha, \beta)$ est la transformée de I_1, celle de I_2 est évidemment $g_1 \exp(-ik\alpha f\theta)$ puisque les deux répartitions d'intensité se déduisent l'une de l'autre par une simple translation. Nous pouvons alors écrire l'intensité reçue dans la direction α, β (qui en pratique serait reçue à distance finie au foyer d'une lentille) comme :

$$I(\alpha, \beta) = 2|g_1|^2 (1 + \cos k\alpha f\theta) \ . \tag{3.5}$$

Le terme en $|g_1|^2$ représente la transformée de Fourier de la tache de speckle. C'est donc une tache irrégulière dont l'extension maximale $1/k(D/\lambda f) \simeq D/f$ correspond à la taille minimale des speckles. Les structures intérieures ont une extension caractéristique $\Delta x_1/f$ correspondant à l'extension totale de la tache de speckle. L'intensité diffractée totale superpose à cette tache aléatoire une modulation certaine à la fréquence $kf\theta/2$. L'examen de ces franges modulant la tache permet évidemment de mesurer la distance angulaire entre les deux étoiles. Il faut pour cela qu'on ait au moins une frange sur l'extension de la tache, ce qui s'exprime par:

$$\frac{D}{f} > \frac{1}{f\theta k} \ , \tag{3.6}$$

ou encore

$$\theta > \frac{\lambda}{D} \ . \tag{3.7}$$

On peut donc séparer les deux étoiles dès que leur distance angulaire est plus grande que le pouvoir de résolution théorique du télescope en l'absence de toute perturbation atmosphérique. On a restitué à l'instrument tout son pouvoir de résolution!

La méthode de Labeyrie peut maintenant être réalisée de manière numérique. En enregistrant un ensemble de taches de speckle, on peut, par des techniques de reconstruction d'image, restituer une véritable image bidimensionnelle. Elle ne peut toutefois s'appliquer qu'à des objets relativement lumineux. On doit en effet enregistrer la tache de speckle en un temps court par rapport aux constantes de temps des fluctuations de l'atmosphère. En pratique, cela impose des temps de pose entre la milliseconde et la dizaine de millisecondes. Seules des étoiles relativement proches sont assez lumineuses. L'optique adaptative, elle, ne souffre pas de cette limite et permet d'obtenir le pouvoir de résolution théorique, même avec des objets très faibles.

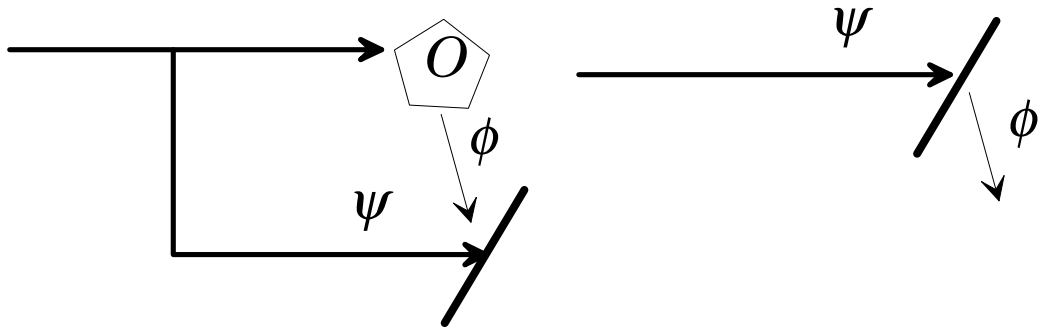

Figure 3.3: Principe de l'enregistrement et de la restitution d'une image holographique. A gauche, enregistrement. La plaque photographique enregistre l'interférogramme entre une onde plane de référence, dérivée de la source cohérente éclairant l'objet, et l'onde diffusée. A gauche, restitution. La plaque photographique, éclairée par l'onde de référence, diffracte entre autres une onde identique à l'onde provenant de l'objet: on obtient une image en relief.

3.3 Holographie

Une des applications les plus connues de la diffraction est sans conteste l'holographie qui est maintenant entrée dans le domaine des applications grand public. L'objet essentiel de l'holographie est de restituer une vision en relief de l'objet enregistré. Il faut pour cela être capable de stocker sur une plaque photographique non seulement le module de l'onde reçue (ce que fait la photographie ordinaire) mais aussi sa phase. La vision en relief n'est complètement restituée que si l'on dispose de tout le front d'onde de l'onde d'origine. Nous commencerons par exposer brièvement, sans entrer dans le détail des calculs, le principe de l'holographie, inventée par Gabor il y a plus d'un demi siècle, mais qui n'est guère entrée dans les applications pratiques qu'avec l'invention du laser. Nous donnerons ensuite brièvement quelques applications non ludiques de l'holographie. Nous montrerons enfin comment l'holographie peut être réalisée en temps réel avec un milieu optique non linéaire.

3.3.1 Principe

Le but de l'opération est de restituer toute l'information de l'onde ϕ rayonnée par un objet éclairé par une source (nous ne considérerons ici que des rayonnements cohérents monochromatiques). Pour cela, on enregistre sur une plaque photographique une interférence entre l'onde ϕ émise par l'objet et une onde plane cohérente Ψ (dérivée du faisceau qui éclaire l'objet et provenant, en pratique, du même laser). Le principe de l'enregistrement de l'hologramme est représenté sur la partie gauche de la figure 3.3. L'amplitude reçue en un point de la plaque photographique est alors $\phi + \Psi$. Après développement, la plaque photographique devient un objet dont la transparence en un point est:

$$T(x, y) = |\phi|^2 + |\Psi|^2 + \phi\Psi^* + \phi^*\Psi \ . \tag{3.8}$$

Les deux premiers termes décrivent une information d'intensité qui ne présente guère d'intérêt. L'enregistrement holographique est en fait contenu dans les deux derniers termes, termes d'interférences et dépendant donc de la phase de l'onde diffusée.

La restitution de l'hologramme consiste à éclairer la plaque photographique avec la même onde de référence Ψ qu'à l'enregistrement. L'onde transmise a alors la forme $T\Psi$ qu'on peut écrire facilement comme:

$$(|\phi|^2 + |\Psi|^2)\Psi + |\Psi|^2\phi + \Psi^2\phi^* \tag{3.9}$$

Le premier terme est essentiellement une onde plane identique à l'onde de référence, multipliée par une fonction de transparence modulée qui introduira un fond de diffraction ne contenant guère d'informations. Le second terme est le terme essentiel. Il est en effet, à des facteurs d'amplitude

près, identique à l'onde diffusée par l'objet. On "voit" donc sortir de la plaque photographique, superposée à une lumière diffuse due aux autres termes, l'onde diffusée par l'objet. En un mot, on observe en arrière de la plaque une image virtuelle de l'objet.

Le dernier terme est important quant on éclaire la plaque holographique avec une onde Ψ^*, c'est à dire une onde plane se propageant dans la direction inverse de Ψ. Superposée au fond, on a en effet dans ce cas une onde ϕ^*, conjuguée de phase de l'onde émise par l'objet. Au lieu d'une onde divergente provenant de l'objet, on a une onde convergeant vers le symétrique de l'objet par rapport à la plaque. En un mot, on observe une image réelle de l'objet en avant de la plaque. Ces "hologrammes avant" sont bien sûr beaucoup plus spectaculaires que les hologrammes arrière traditionnels, parce que l'objet semble sortir de la plaque et flotter dans l'espace. Une visite dans une exposition d'hologrammes s'impose à ce point. Insistons encore sur le fait que l'hologramme restitue toute l'information. L'hologramme d'un objet sous une loupe, par exemple, contient l'objet, la loupe, mais aussi l'image de l'objet agrandi par la loupe!

Nous pouvons, à partir de cette simple discussion, dégager les conditions expérimentales pour l'obtention d'un hologramme. D'abord, il est absolument crucial de disposer d'ondes présentant une stabilité de phase parfaite pendant le temps d'enregistrement. On peut réaliser une stabilité approchée, comme le faisait Gabor, avec des sources ordinaires, convenablement filtrées spatialement et en fréquence. Mais leur intensité devient très faible, les temps de pose prohibitifs et on ne peut enregistrer que des hologrammes de petits objets très lumineux. La méthode ne présente d'intérêt pratique qu'avec une source laser, très intense et parfaitement cohérente temporellement. Il faut cependant que les distances relatives entre l'objet, le laser et la plaque photographique restent constantes à beaucoup mieux qu'une longueur d'onde pendant le temps de pose. Il faut donc un montage interférométriquement stable, ce qui rend plus difficile l'enregistrement d'objets mobiles ou la cinématographie holographique. Enfin, il faut que la plaque photographique ait un grain très fin pour enregistrer fidèlement un figure d'interférences dont le pas est évidemment de l'ordre de la longueur d'onde. Un grain aussi fin implique une sensibilité relativement basse et des temps de pose longs.

Enfin, un point essentiel est de restituer l'hologramme avec une onde cohérente. Dans les premiers temps de l'holographie, on devait utiliser le laser d'enregistrement, ce qui rendait la lecture complexe. On a depuis mis au point des hologrammes visibles en lumière naturelle, tels ceux qui servent de cachet d'authentification sur les cartes de crédit. L'idée est tirée en fait du principe de la photographie Lippman en couleurs. Dans cette technique, on enregistre une image photographique avec une émulsion épaisse préparée sur un miroir. Une longueur d'onde donnée crée alors une onde stationnaire. Après développement de l'émulsion, on obtient des plans métalliques (argent) distants d'une demi–longueur d'onde du rayonnement incident (en incidence normale). Ces plans constituent une série d'interféromètres de Fabry Perot qui ne réfléchissent efficacement que les longueurs d'onde très voisine de celle d'enregistrement. Lippman pouvait ainsi réaliser des photographies de couleurs très réalistes avec une émulsion monochrome. Contrairement aux photographies couleurs basées sur des pigments, celles de Lippman sont parfaitement stables à la lumière et conservent leurs couleurs après près d'un siècle. Dans le domaine de l'holographie, le même principe peut être utilisé pour ne réfléchir que les longueurs d'onde voisines de la longueur d'onde incidente et fournir une cohérence suffisante pour lire l'hologramme. Cela explique aussi que les couleurs interférentielles observées sur ces hologrammes en lumière naturelle changent rapidement avec l'angle d'incidence.

Notons enfin que l'information sur l'objet est délocalisée sur toute la surface de l'hologramme. Une portion quelconque de l'hologramme peut être utilisée pour la reconstruction. Elle régénère en effet toute l'amplitude diffusée par l'objet et pas, comme on pourrait s'y attendre intuitivement, l'amplitude diffusée par une partie de l'objet seulement. On y perd seulement sur le champ de vision (l'objet apparaît vu à travers la plaque holographique) et sur la résolution spatiale de la reconstruction (un peu de la même manière qu'on perd sur la résolution spectrale d'un réseau en diminuant le nombre de traits).

3.3.2 Applications

Au delà d'aspects ludiques ou artistiques, l'holographie est un précieux moyen d'observation. Une application industrielle particulièrement importante est la visualisation des vibrations des machines en mouvement. Supposons qu'on prenne, à un instant t, un hologramme instantané d'une machine en fonctionnement. On peut réaliser ce genre d'hologramme en utilisant un laser en impulsions brèves. La durée d'impulsion, en général de l'ordre de la dizaine de nanosecondes, est assez brève pour que tous les points de la machine puissent être considérés comme immobiles, assez longue cependant pour que le faisceau soit cohérent (la longueur de cohérence est alors de quelques mètres[2]). L'onde enregistrée sur la plaque holographique sera notée ϕ. Pour simplifier notre analyse qualitative, nous considérerons d'abord que la machine est réduite à un seul point en mouvement. A l'instant t, nous avons donc réalisé l'hologramme d'un point à la position $\mathbf{r}$. A un instant ultérieur t', on enregistre, sur la même plaque, l'onde ϕ' correspondant à la nouvelle position $\mathbf{r}'$ du point. On développe (ou on traite numériquement dans les versions modernes). La transmission de la plaque contient alors un terme en $\Psi^*(\phi + \phi')$. L'onde restituée est alors $\phi + \phi'$. On aurait la même image en éclairant, à un seul instant, deux points sources situés aux positions $\mathbf{r}$ et $\mathbf{r}'$ avec une seule source cohérente. La relecture de l'hologramme doit donc fournir des interférences entre ces deux points sources cohérents. Quand on tient compte de tous les points de l'objet en mouvement, on observe sur l'image reconstituée des franges d'interférences qui révèlent les déplacements de l'objet entre les instants t et t'. Ces franges sont analogues aux franges d'égale épaisseur des lames minces d'air. Elles sont localisées à la surface de l'objet. Le simple examen de l'image holographique révèle les mouvements avec une résolution égale à la longueur d'onde optique. En synchronisant les impulsions laser sur le fonctionnement de l'appareil, on peut réaliser une analyse stroboscopique des vibrations. Cette méthode n'est pas plus sensible que les méthodes interférométriques standard mais elle permet une vue d'ensemble du mouvement de la pièce.

Une autre application importante de l'holographie est la réalisation de composants optiques. Imaginons, par exemple, qu'on prenne l'hologramme d'un point source. Eclairé par l'onde de référence, une onde plane, il fournit une image virtuelle du point source. L'onde plane incidente se transforme en une onde sphérique divergente. L'hologramme se comporte donc comme une lentille divergente. Si, au contraire, on réalise un hologramme avant du point source, l'onde plane incidente devient une onde sphérique convergente et l'hologramme est une lentille convergente.

Réaliser ainsi une simple lentille ne présente guère d'intérêt dans le domaine optique. L'intensité de l'onde reconstituée est en effet notablement inférieure à l'intensité de l'onde transmise. L'intérêt est plus important quand on entre dans le domaine des rayons X. Tous les matériaux dans ce domaine ont en effet un indice de réfraction extrêmement voisin de un et il est impossible de réaliser des composants d'optique réfractive. On peut réaliser des miroirs avec des matériaux cristallins mais ils ne fonctionnent qu'en incidence rasante et sont extrêmement coûteux. En revanche, on peut aisément fabriquer, par des techniques de microlithographie, des structures absorbantes avec des détails extrêmement fins. On peut donc synthétiser l'hologramme qui se comportera comme une lentille. Un domaine nouveau de microscopie X a pu apparaître ainsi avec les optiques holographiques et les sources intenses de rayonnement synchrotron. En utilisant l'analogie formelle entre équations de Maxwell et équation de Schrödinger, on peut aussi préparer des hologrammes pour des ondes de matière. Il est ainsi possible de focaliser des ondes de matière ou de réaliser des images avec des atomes refroidis par laser. Avec des températures ordinaires, les longueurs d'onde de de Broglie des atomes sont si petites (de l'ordre du picomètre) qu'on ne peut réaliser les structures correspondantes. Avec des atomes refroidis, en revanche, la longueur d'onde de de Broglie est de l'ordre d'une longueur d'onde optique. On peut alors aisément réaliser des structures holographiques. Les très belles expériences de Shimizu (Tokyo) ont démontré la faisabilité de cette technique qui pourrait être employée dans le domaine de

[2]La longueur de cohérence, dans ce cas très simple, est au plus égale au produit de la vitesse de la lumière par la durée de l'impulsion. Il s'agit de la différence de marche maximale qui donne lieu à des interférences visibles dans un interféromètre de Michelson

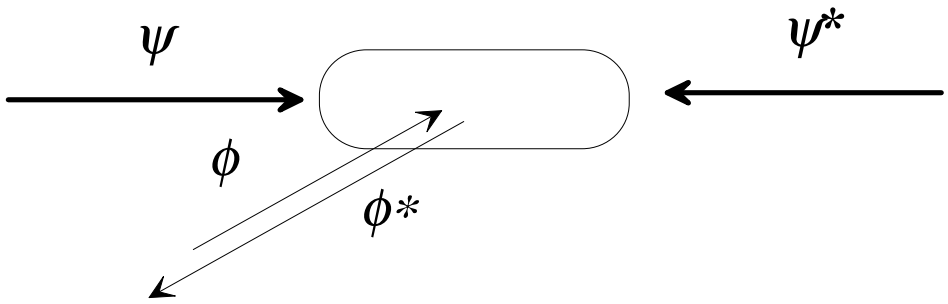

Figure 3.4: Pincipe d'une expérience de conjugaison de phase optique, qui se ramène à une holographie en temps réel.

la microlithographie pour des dépôts très bien contrôlés.

Les composants holographiques sont utiles également pour réaliser des fonctions optiques non classiques. Le balayage du faisceau laser sur le tambour photosensible des imprimantes laser ou du faisceau des lecteurs de codes barres, par exemple, est réalisé par un miroir holographique mis en rotation. Le coût de production de masse de ces objets est infiniment inférieur à celui des miroirs prismatiques initialement utilisés.

3.3.3 Holographie et conjugaison de phase

Telle que nous l'avons décrite jusqu'ici, l'holographie procède en deux temps: enregistrement de la figure d'interférences et restitution ultérieure. Il est possible de réaliser les deux étapes simultanément si on peut disposer d'un matériau dont l'indice ou l'absorption dépende de l'intensité en chaque point. C'est effectivement le cas avec les matériaux optiques non linéaires. Considérons le dispositif représenté sur la figure 3.4. Un cristal spécial (on utilise souvent du Niobiate de Lithium ou une de ses variantes) est éclairé par deux ondes planes intenses contra–propageantes, Ψ et Ψ^*. On envoie, dans une autre direction, une onde ϕ qui joue le rôle de l'onde diffusée en holographie. Les ondes Ψ et Ψ^* vont bien sûr jouer le rôle des ondes de référence et de reconstruction. Le matériau non linéaire acquiert, en plus de la polarisation linéaire qui décrit simplement l'indice, une densité de polarisation P proportionnelle au produit de trois termes qui sont les champs incidents ou leurs complexes conjugués. Pour décrire vectoriellement ce processus, il faudrait bien sûr une polarisabilité tensorielle de rang 4. Pour une discussion qualitative, nous nous contenterons de voir que cette non–linéarité confère au milieu un indice de réfraction n qui, en plus d'un terme constant correspondant à la réponse linéaire (voir partie 5), présente un terme non linéaire proportionnel à l'intensité. Parmi tous les termes apparaissant dans le développement de cette non–linéarité sur les amplitudes des différentes ondes, on aura ceux correspondant à $|\phi + \Psi|^2$. Sur le milieu, s'inscrit donc un réseau de modulation d'indice à trois dimensions, qui est précisément l'hologramme de l'onde ϕ par l'onde de référence Ψ. La "lecture" simultanée de cet hologramme par l'onde Ψ^* va fournir une onde ϕ^* qui est précisément la conjuguée complexe de l'onde ϕ. Ce milieu réalise donc la conjugaison de phase de l'onde incidente, ce qui ne peut s'effectuer par aucun dispositif d'optique linéaire.

Nous avons ici vu de manière globale le principe de la génération. On peut aussi le comprendre en termes holographiques. Le mélange de Ψ et de ϕ crée dans le milieu non linéaire une répartition d'intensité qui n'est autre que l'hologramme de l'onde ϕ avec la référence Ψ. En raison de son caractère non linéaire, le milieu acquiert un indice de réfraction modulé spatialement par cet hologramme. L'onde Ψ^* est diffractée par cette répartition d'indice. En fait, elle vient simplement lire l'hologramme avant et recrée l'onde ϕ^*.

Un tel miroir à conjugaison de phase présente des propriétés optiques extraordinaires. Imaginons qu'on place en face de lui un point source. Il crée une onde sphérique divergente. En se réfléchissant sur le miroir, elle devient une onde sphérique convergente qui revient se focaliser exactement sur le point

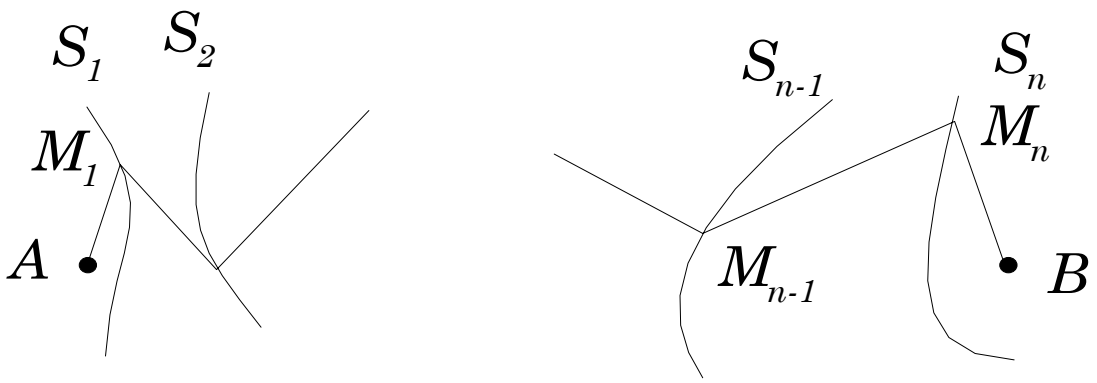

Figure 3.5: Passage à la limite géométrique. L'amplitude reçue du point source A au point B peut se mettre sous la forme d'une somme de termes associés à toutes les lignes brisées passant de A à B ou de tout chemin à la limite continue. Seuls les chemins proches de la trajectoire prévue par l'optique géométrique contribuent notablement.

source. Quand on sait que rien n'empêche un miroir à conjugaison de phase d'avoir un coefficient de réflexion supérieur à un (l'énergie nécessaire étant prise dans les ondes intenses Ψ et Ψ^*), on comprendra l'intérêt potentiel de ce dispositif. On peut ainsi par exemple réaliser un laser constitué d'un miroir à conjugaison de phase et de n'importe quel objet vaguement réfléchissant passant à proximité. La conjugaison de phase assure le retour parfait de l'onde sur l'objet réfléchissant et les conditions de phase correctes pour créer une cavité.

On peut aussi utiliser la conjugaison de phase pour optimiser les communications. Dans un matériau hétérogène comme l'atmosphère, la propagation distord les fronts d'onde. Un point source donne donc un front d'onde loin de la forme sphérique idéale. Si on renvoie cette onde par un miroir à conjugaison de phase, on change le signe des défauts de phase. Le retour inverse dans le milieu (s'il n'a pas notablement varié pendant le temps d'un aller et retour) annule exactement ces défauts et l'onde de retour se focalise exactement sur le point source comme s'il n'y avait pas eu de défauts.

3.4 Limite de l'optique géométrique

La diffraction peut nous permettre de mieux comprendre comment se fait le passage de l'électromagnétisme de Maxwell à la notion de rayons lumineux. Ce point sera discuté très en détail dans l'appendice 4 à cette partie mais nous allons donner ici l'essentiel des arguments qualitatifs nécessaires. Nous venons de voir qu'une onde plane est très peu affectée, très peu diffractée, si elle ne passe qu'à travers des ouvertures ayant une taille notablement plus grande que λ. Il est donc possible de réaliser un pinceau lumineux rectiligne, dans un milieu homogène, dont l'extension soit grande devant λ et néanmoins très petite à l'échelle macroscopique qui est celle de la plupart des composants optiques. Nous venons de retrouver la notion de rayon lumineux. Comme les lois de changement de vecteur d'onde au passage entre deux matériaux diélectriques sont précisément les lois de Descartes (on consultera à ce sujet la fin de la cinquième partie), ces rayons lumineux obéissent précisément aux lois de l'optique géométrique.

Nous allons préciser un peu ces notions en montrant qualitativement, à partir de la formule de Kirchhoff, qu'on peut exprimer le champ reçu en un point comme une somme d'amplitudes associées à différents chemins possibles connectant ce point au point source. Nous montrerons alors que seuls les chemins situés dans un tube d'une extension transverse de quelques λ autour de la trajectoire prédite par l'optique géométrique contribuent effectivement à la somme. La situation que nous décrivons est représentée schématiquement sur la figure 3.5.

On cherche à calculer l'amplitude reçue en B, la source étant située en A. Nous supposerons d'abord que le milieu entre A et B est homogène. Formellement, en utilisant la formule (2.24), on peut exprimer le champ (au sens d'une composante du potentiel) en B en fonction du champ sur une

surface S_n entourant complètement B et passant à très faible distance au niveau de la ligne droite qui joint A à B. Nous n'avons pas ici d'approximation liée au principe de Huyghens, puisqu'il n'y a pas de diaphragme. Pour une surface S_n suffisamment régulière, les approximations assimilant l'onde incidente à une onde plane (au moins localement) et la surface à un plan (localement là aussi) sont sans doute valables. On aura donc:

$$\phi_B \simeq \int_{S_n} \chi_n \frac{\phi_n(\mathbf{r}_n)\exp(ik|\mathbf{r}_n - \mathbf{r}_B|)}{|\mathbf{r}_n - \mathbf{r}_B|}\, dS_n \ , \tag{3.10}$$

où $\mathbf{r}_n$ est la position courante sur S_n, ϕ_n l'amplitude en ce point et χ_n le facteur d'inclinaison géométrique.

L'amplitude ϕ_n peut elle aussi s'exprimer comme l'intégrale sur une surface S_{n-1}, proche de S_n. Et ainsi de suite, jusqu'à une surface S_1 entourant immédiatement le point source A. On aura finalement ϕ comme une intégrale sur toutes les surfaces simultanément, c'est à dire, aussi bien, une intégrale sur toutes les lignes brisées qui passent de A à B en s'appuyant sur les surfaces intermédiaires. Le terme à intégrer contient des termes d'amplitudes, lentement variables et de peu d'importance comme on le verra. Le terme important est le terme de phase qui s'écrit simplement:

$$\exp[ik(AM_1 + M_1M_2 + \cdots + M_nB)] = e^{ikL} \ , \tag{3.11}$$

où les M_i sont les points de la ligne brisée sur les surfaces intermédiaires et L est la longueur totale de la ligne brisée.

Les phases associées à deux lignes brisées différentes sont complètement différentes dès que ces lignes s'écartent l'une de l'autre d'une petite fraction de λ. Les amplitudes étant comparables, les contributions à l'amplitude de toutes ces lignes sont négligeables. Toutes, sauf celles qui se trouvent au voisinage d'une trajectoire qui rend extrémale la longueur du trajet. Il s'agit en l'occurrence d'un minimum et de la simple ligne droite entre A et B. Toutes les trajectoires ne s'écartant que de quelques λ par rapport à cette ligne contribuent de façon équivalente à la somme. Toutes les trajectoires passant plus loin ont une influence négligeable. On peut donc insérer un diaphragme de taille grande devant λ au voisinage de cette ligne droite sans modifier la propagation: on retrouve bien, de façon moins qualitative, la notion de rayon lumineux.

Notons que le raisonnement que nous faisons ici dans le domaine optique peut s'appliquer, avec des modifications mineures, à la mécanique quantique. On trouve qu'une amplitude de transition est la somme d'amplitudes associées à tous les chemins possibles entre l'état initial et l'état final. Dans une limite semi–classique, les effets quantiques sont presque négligeables et seuls contribuent les chemins situés à quelques longueurs d'onde de de Broglie de la trajectoire classique.

Le raisonnement précédent peut être aisément modifié pour prendre en compte les inhomogénéités de l'indice de réfraction du milieu. Le vecteur d'onde local étant proportionnel à l'indice, on trouve que le facteur de phase peut s'écrire dans ce cas:

$$\exp[ik_0(n_1AM_1 + n_2M_1M_2 + \cdots + n_nM_nB)] = e^{ik_0\int ndl} \ , \tag{3.12}$$

où n_i est l'indice au voisinage du point M_i. On trouve donc que ne contribuent notablement que les trajectoires situées au voisinage immédiat de celles qui extrémalisent l'intégrale $\int ndl$, c'est à dire le temps de parcours. En un mot, nous retrouvons le principe de Fermat qui est à la base de l'optique géométrique. Encore une fois, ces raisonnements ne sont que qualitatifs. Nous les rendrons beaucoup plus précis dans l'appendice 4.

Appendice 1

Jauges

Nous reviendrons très brièvement dans cet appendice sur les problèmes de choix de jauge. Nous savons bien que les équations de Maxwell et donc toutes les prédictions physiques sont invariantes dans une modification des potentiels s'écrivant:

$$\mathcal{A}'^{\mu} = \mathcal{A}^{\mu} + \partial^{\mu}\phi \tag{1.1}$$

en notations quadridimensionnelles ou

$$\mathbf{A}' = \mathbf{A} + \boldsymbol{\nabla}\phi \tag{1.2}$$

$$V' = V - \frac{\partial\phi}{\partial t} \tag{1.3}$$

en notations standard, ϕ étant, dans ces expressions, une fonction arbitraire de l'espace et du temps. On peut utiliser cette liberté de jauge pour tenter de simplifier la forme des équations décrivant les potentiels.

Le choix le plus répandu, celui que nous avons déjà largement utilisé, est celui de la jauge de Lorentz qui s'écrit:

$$\partial_{\mu}\mathcal{A}^{\mu} = 0 \tag{1.4}$$

ou encore

$$\boldsymbol{\nabla}\cdot\mathbf{A} + \frac{1}{c^2}\frac{\partial V}{\partial t} = 0 \ . \tag{1.5}$$

Cette jauge a plusieurs mérites. D'abord, elle est réalisable. Partant de potentiels $\mathcal{A}'^{\mu}$ vérifiant une condition de jauge arbitraire, on peut leur rajouter la 4–divergence d'une fonction ϕ pour qu'ils vérifient la jauge de Lorentz. Il suffit pour cela de résoudre:

$$\partial_{\mu}\partial^{\mu}\phi = \partial_{\mu}\mathcal{A}'^{\mu} \ , \tag{1.6}$$

ce qui est possible avec des conditions aux limites ordinaires (tous les potentiels nuls à l'infini dans l'espace temps). Ensuite, la jauge de Lorentz est explicitement covariante : elle s'écrit comme la nullité d'un 4–scalaire. Enfin, et ce n'est pas son moindre mérite, elle conduit à des équations aux potentiels extrêmement simples, puisqu'elles découplent le potentiel vecteur dont la seule source est le courant et le potentiel scalaire dont la seule source est la densité de charges. Nous avons vu dans le premier chapitre que ces équations admettaient la solution explicite en termes de potentiels retardés.

Si la jauge de Lorentz présente des avantages importants, il en est une autre qui est largement utilisée, en particulier dans le domaine de la physique atomique. C'est la jauge de la magnétostatique, ou jauge de Coulomb. Elle s'écrit

$$\boldsymbol{\nabla}\cdot\mathbf{A} = 0 \ . \tag{1.7}$$

Manifestement, elle n'est pas covariante. Si elle est satisfaite dans un référentiel, elle ne le sera plus après une transformation de Lorentz. Cette jauge n'a donc de sens que dans un référentiel donné,

dans lequel nous nous placerons désormais. Nous devons maintenant vérifier que cette jauge peut être satisfaite. Supposons qu'on ait des potentiels $\mathbf{A}'$ et V' qui ne satisfont pas à la jauge de Coulomb. Les potentiels issus d'une transformation de jauge y satisferont à condition que $\Delta\phi = -\boldsymbol{\nabla}\cdot\mathbf{A}'$. Cette équation est l'équation de Poisson de l'électrostatique qui admet une solution unique avec des conditions aux limites convenables. On peut donc toujours choisir de travailler en jauge de Coulomb.

Ecrivons maintenant les équations aux potentiels dans cette jauge. En raison de la jauge, $\boldsymbol{\nabla}\cdot\mathbf{E} = -\Delta V$ et l'équation de V s'écrit simplement:

$$\Delta V = -\frac{\rho}{\epsilon_0} \ . \tag{1.8}$$

C'est tout simplement l'équation de l'électrostatique. En jauge de Coulomb, le potentiel scalaire se calcule comme en électrostatique. C'est donc un potentiel instantané, en ce sens que le potentiel en $\mathbf{r}, t$ dépend des sources en $\mathbf{r}_1$ au même instant:

$$V = \frac{1}{4\pi\epsilon_0}\int \frac{\rho(\mathbf{r}_1, t)}{|\mathbf{r}-\mathbf{r}_1|}\, d^3\mathbf{r}_1 \ . \tag{1.9}$$

Bien sûr, l'électromagnétisme ne violera pas pour autant la causalité relativiste. Les champs électriques, eux, ne dépendront que des sources aux instants retardés. Clairement, tout l'intérêt de la jauge de Coulomb réside dans cette simplicité du potentiel scalaire. En physique atomique, par exemple, on traite le couplage d'un atome, en liaison coulombienne, avec un champ oscillant. Tout naturellement, la jauge de Coulomb séparera dans la dynamique du système le champ statique liant l'électron, représenté par le simple potentiel scalaire, du champ oscillant incident.

L'équation au potentiel vecteur est en revanche plus complexe. En écrivant que $\boldsymbol{\nabla}\times\mathbf{B} = \boldsymbol{\nabla}\times(\boldsymbol{\nabla}\times\mathbf{A}) = \mu_0\mathbf{j} + (1/c^2)\partial\mathbf{E}/\partial t$ et en substituant l'équation donnant le champ électrique en fonction des potentiels:

$$\Box\,\mathbf{A} = -\mu_0\mathbf{j} + \frac{1}{c^2}\frac{\partial\boldsymbol{\nabla}V}{\partial t} \ . \tag{1.10}$$

Cette équation est beaucoup plus complexe que l'équation au d'alembertien obtenue en jauge de Lorentz. Elle couple effectivement le potentiel vecteur au gradient du potentiel scalaire. En un mot, la résolution d'un problème d'électromagnétisme en jauge de Coulomb revient à d'abord trouver le potentiel scalaire électrostatique, pour ensuite l'utiliser comme terme source dans le potentiel vecteur. Nous verrons dans le prochain appendice comment on peut aborder le problème en séparant variables longitudinales et transverses du champ.

Appendice 2

Espace réciproque

Nous utiliserons dans cet appendice la transformation de Fourier sur les variables d'espace (nous garderons le temps à part pour conserver sous forme directe les propriétés dynamiques) pour transposer les équations de Maxwell dans l'espace réciproque. Après avoir défini les transformations mises en jeu, nous définirons dans le premier paragraphe les champs transverses et longitudinaux. Nous montrerons, dans le second paragraphe, que la jauge de Coulomb introduite dans l'appendice précédent permet de séparer très naturellement la dynamique des champs longitudinaux, purement électrostatiques, de celle des champs transverses. Nous nous pencherons alors sur la dynamique des champs transverses dans l'espace libre et nous explorerons l'analogie formelle entre l'électromagnétisme et la dynamique de l'oscillateur harmonique en introduisant les variables normales du champ. Nous montrerons enfin, dans un dernier paragraphe, le lien entre cette approche et la décomposition en ondes planes introduite dans les classes élémentaires.

Au delà d'une simple curiosité, cet appendice est en fait une introduction directe à la quantification du champ. C'est en effet en termes de variables normales ou de collection d'oscillateurs harmoniques qu'on quantifie le champ. Nous n'aborderons pas du tout ici cette quantification dont nous ne ferons que planter les jalons classiques. Un exposé très clair en est donné dans "Introduction à l'électrodynamique quantique" de Cohen, Dupont-Roc et Grynberg.

2.1 Espace réciproque

Nous allons simplement écrire dans ce paragraphe les équations de l'électromagnétisme (équations de Maxwell et équations aux potentiels) en termes de transformées de Fourier spatiales (l'approche est donc un peu différente de ce que nous avions fait pour établir l'expression de la fonction de Green). Les opérateurs de l'analyse vectorielle devenant de simples vecteurs, l'écriture de ces équations sera bien sûr beaucoup plus simple que dans l'espace réel.

2.1.1 Transformation de Fourier

Nous transformerons les quantités scalaires et vectorielles. Le champ électrique s'écrira par exemple:

$$\mathbf{E}(\mathbf{r},t) = \frac{1}{(2\pi)^{3/2}} \int d^3\mathbf{k}\, \boldsymbol{\mathcal{E}}(\mathbf{k},t) e^{i\mathbf{k}\cdot\mathbf{r}} \ . \tag{2.1}$$

La transformation inverse s'écrit:

$$\boldsymbol{\mathcal{E}}(\mathbf{k},t) = \frac{1}{(2\pi)^{3/2}} \int d^3\mathbf{r}\, \mathbf{E}(\mathbf{r},t) e^{-i\mathbf{k}\cdot\mathbf{r}} \ . \tag{2.2}$$

Le champ électrique étant réel, il vérifie:

$$\mathbf{E}^*(\mathbf{r},t) = \mathbf{E}(\mathbf{r},t) \ . \tag{2.3}$$

On a donc:

$$\int \boldsymbol{\mathcal{E}}^*(\mathbf{k},t)e^{-i\mathbf{k}\cdot\mathbf{r}}d^3\mathbf{k} = \int \boldsymbol{\mathcal{E}}(\mathbf{k},t)e^{i\mathbf{k}\cdot\mathbf{r}}d^3\mathbf{k} = \int \boldsymbol{\mathcal{E}}(-\mathbf{k},t)e^{-i\mathbf{k}\cdot\mathbf{r}}d^3\mathbf{k} \ . \tag{2.4}$$

Si cette propriété est vraie pour tout $\mathbf{r}$, alors:

$$\boldsymbol{\mathcal{E}}(-\mathbf{k},t) = \boldsymbol{\mathcal{E}}^*(\mathbf{k},t) \ . \tag{2.5}$$

Nous ferons la même opération sur le champ magnétique (transformée de Fourier $\boldsymbol{\mathcal{B}}$), les potentiels $\boldsymbol{\mathcal{A}}$ et $\mathcal{V}$ et les sources ($\tilde{\rho}$ et $\tilde{\mathbf{j}}$). En termes de ces variables, les équations de Maxwell s'écrivent simplement:

$$i\mathbf{k}\times\boldsymbol{\mathcal{E}} = -\dot{\boldsymbol{\mathcal{B}}} \tag{2.6}$$
$$i\mathbf{k}\times\boldsymbol{\mathcal{B}} = \mu_0\tilde{\mathbf{j}} + \dot{\boldsymbol{\mathcal{E}}}/c^2 \tag{2.7}$$
$$i\mathbf{k}\cdot\boldsymbol{\mathcal{E}} = \tilde{\rho}/\epsilon_0 \tag{2.8}$$
$$i\mathbf{k}\cdot\boldsymbol{\mathcal{B}} = 0 \tag{2.9}$$

(nous avons indiqué par un point les dérivations temporelles dans l'espace réciproque pour alléger les notations). L'équation de conservation de la charge devient:

$$i\mathbf{k}\cdot\tilde{\mathbf{j}} + \dot{\tilde{\rho}} = 0 \ , \tag{2.10}$$

les liens entre champs et potentiels s'écrivent:

$$\boldsymbol{\mathcal{B}} = i\mathbf{k}\times\boldsymbol{\mathcal{A}} \tag{2.11}$$
$$\boldsymbol{\mathcal{E}} = -i\mathbf{k}\mathcal{V} - \dot{\boldsymbol{\mathcal{A}}} \ . \tag{2.12}$$

Enfin les jauges de Coulomb et de Lorentz s'écrivent respectivement:

$$i\mathbf{k}\cdot\boldsymbol{\mathcal{A}} = 0 \tag{2.13}$$

et

$$i\mathbf{k}\cdot\boldsymbol{\mathcal{A}} + \dot{\mathcal{V}}/c^2 = 0 \ . \tag{2.14}$$

Le lecteur pourra à titre d'exercice transposer ainsi d'autres équations importantes de l'électromagnétisme.

2.1.2 Champs longitudinaux et transverses

Les équations de Maxwell dans l'espace réciproque font intervenir des produits scalaires et vectoriels avec le vecteur d'onde. Il est donc logique et probablement fructueux de décomposer les champs et potentiels vectoriels en leurs composantes longitudinales, parallèles à $\mathbf{k}$, et transverses, perpendiculaires à $\mathbf{k}$. Introduisant le vecteur unitaire $\boldsymbol{\kappa} = \mathbf{k}/k$, nous définirons par exemple la partie longitudinale du champ électrique dans l'espace réciproque par:

$$\boldsymbol{\mathcal{E}}_{\parallel} = (\boldsymbol{\kappa}\cdot\boldsymbol{\mathcal{E}})\boldsymbol{\kappa} \tag{2.15}$$

et la partie tranverse par:

$$\boldsymbol{\mathcal{E}}_{\perp} = \boldsymbol{\kappa}\times(\boldsymbol{\mathcal{E}}\times\boldsymbol{\kappa}) \ . \tag{2.16}$$

Le champ électrique total dans l'espace réciproque est bien sûr la somme de ses composantes longitudinales et transverses:

$$\boldsymbol{\mathcal{E}}(\mathbf{k},t) = \boldsymbol{\mathcal{E}}_{\parallel} + \boldsymbol{\mathcal{E}}_{\perp} \ . \tag{2.17}$$

La même décomposition peut s'effectuer pour le champ magnétique et le potentiel vecteur.

La linéarité de la transformation de Fourier nous permet alors d'écrire les champs dans l'espace réel comme une somme de deux composantes, transformées de Fourier inverses des champs transverses et longitudinaux dans l'espace réciproque. On aura ainsi:

$$\mathbf{E}(\mathbf{r},t) = \mathbf{E}_{\parallel}(\mathbf{r},t) + \mathbf{E}_{\perp}(\mathbf{r},t) \ , \tag{2.18}$$

où $\mathbf{E}_{\parallel}$ est la transformée de $\boldsymbol{\mathcal{E}}_{\parallel}$ et $\mathbf{E}_{\perp}$ celle de $\boldsymbol{\mathcal{E}}_{\perp}$. Il est évident que les désignations "longitudinal" ou "transverse" ne correspondent à aucune propriété géométrique particulière dans l'espace réel. Seuls les champs dans l'espace réciproque ont une interprétation géométriquement simple. Notons aussi qu'on peut exprimer directement les composantes transverses (et donc, par différence, longitudinales) du champ dans l'espace réel à partir du champ total. La transformation fait intervenir un noyau intégral qu'on trouvera détaillé dans *Introduction à l'électrodynamique quantique* de Cohen. On y constatera que l'expression d'une des composantes du champ transverse fait apparaître toutes les composantes du champ complet.

Les conditions de Jauge s'expriment facilement en fonction des composantes transverses et longitudinales. La jauge de Coulomb, en particulier, devient très simple, puisqu'elle exprime simplement que le potentiel vecteur $\boldsymbol{\mathcal{A}}$ est purement transverse. On a donc:

$$\boldsymbol{\mathcal{A}}_{\parallel} = 0 \ . \tag{2.19}$$

Cette jauge, en dépit de son manque d'invariance relativiste, rend beaucoup plus simple la géométrie des champs dans l'espace réciproque.

Nous nous placerons donc, jusqu'à la fin de cet appendice, en jauge de Coulomb.

Comme $\boldsymbol{\mathcal{A}}$ est transverse, $\mathbf{k} \times \boldsymbol{\mathcal{A}} = \boldsymbol{\mathcal{B}}$ l'est aussi, ainsi que $\partial\boldsymbol{\mathcal{A}}/\partial t$. En revanche, le gradient du potentiel scalaire est longitudinal. On a donc directement:

$$\boldsymbol{\mathcal{E}}_{\perp} = -\dot{\boldsymbol{\mathcal{A}}} \tag{2.20}$$

$$\boldsymbol{\mathcal{E}}_{\parallel} = -i\mathbf{k}\mathcal{V} \ . \tag{2.21}$$

En jauge de Coulomb, le potentiel scalaire se calcule comme en électrostatique. Le champ électrique longitudinal dans l'espace réel $\mathbf{E}_{\parallel}$ est donc simplement le champ électrique calculé comme en électrostatique! On voit ici, de manière plus détaillée que dans l'appendice précédent, tout l'intérêt de la jauge de Coulomb pour traiter un problème de physique atomique. Considérons un atome d'hydrogène soumis à une onde plane et essayons d'écrire les équations du mouvement de l'électron. L'onde incidente n'a pas de potentiel scalaire. Le champ électrique longitudinal est donc entièrement dû au proton. Il se calcule comme en électrostatique et donne un potentiel de liaison Coulombien en $1/r$. L'onde plane, pour sa part, contribue uniquement à la partie transverse des champs. De manière toute naturelle, la jauge de Coulomb sépare la contribution, en général dominante, de l'interaction électrostatique avec le noyau de la contribution du champ incident. Cette dernière peut souvent être considérée comme une perturbation simple par rapport au champ de liaison (ce n'est que dans le domaine des impulsions laser très brèves et très intenses qu'on peut sortir de ce régime et atteindre même des situations où le champ de liaison devient une petite perturbation par rapport au champ laser incident).

2.2 Variables normales

Nous allons dans ce paragraphe souligner l'analogie très forte entre la dynamique des champs électromagnétiques et celle d'une collection d'oscillateurs harmoniques. On avait déjà pu prendre conscience de cette analogie en écrivant la densité de lagrangien du champ électromagnétique, en $E^2 - B^2$, et en remarquant l'analogie formelle avec le lagrangien d'un oscillateur unidimensionnel, en $v^2 - x^2$. Pour ne pas trop alourdir le débat, nous ne discuterons de cette analogie que dans l'espace libre de charges et de courants.

2.2.1 Electromagnétisme dans l'espace libre

Si $\tilde{\rho}$ et $\tilde{\mathbf{j}}$ sont nuls, les équations de Maxwell dans l'espace réciproque se simplifient beaucoup. On a

$$\dot{\boldsymbol{\mathcal{E}}} = ic^2\mathbf{k} \times \boldsymbol{\mathcal{B}} \tag{2.22}$$

et

$$\dot{\boldsymbol{\mathcal{B}}} = -i\mathbf{k} \times \boldsymbol{\mathcal{E}} \tag{2.23}$$

Les champs électriques et magnétiques sont strictement transverses. On en déduit à nouveau que le potentiel scalaire $\mathcal{V}$ est nul.

Nous introduirons, pour décrire le champ, deux nouvelles variables que nous nommerons *variables normales du champ*:

$$\boldsymbol{\alpha}(\mathbf{k},t) = -\frac{i}{2\mathcal{N}}(\boldsymbol{\mathcal{E}} - c\boldsymbol{\kappa} \times \boldsymbol{\mathcal{B}}) \tag{2.24}$$

$$\boldsymbol{\beta}(\mathbf{k},t) = -\frac{i}{2\mathcal{N}}(\boldsymbol{\mathcal{E}} + c\boldsymbol{\kappa} \times \boldsymbol{\mathcal{B}}) \ . \tag{2.25}$$

Ce changement de variables et bien sûr inversible et $\boldsymbol{\alpha}$ et $\boldsymbol{\beta}$ suffisent à déterminer complètement les champs. Dans les expressions précédentes, $\mathcal{N}$ est un facteur de normalisation, pouvant dépendre de $\mathbf{k}$ mais pas de t, que nous ne préciserons pas davantage[1]. Notons également que, par construction, ces variables sont uniquement transverses. La condition de réalité des champs $\boldsymbol{\mathcal{E}}^*(\mathbf{k},t) = \boldsymbol{\mathcal{E}}(-\mathbf{k},t)$ se reflète sur $\boldsymbol{\alpha}$ et $\boldsymbol{\beta}$. On vérifiera immédiatement, par simple substitution, que:

$$\boldsymbol{\beta}^*(\mathbf{k},t) = -\boldsymbol{\alpha}(-\mathbf{k},t) \tag{2.26}$$

$$\boldsymbol{\alpha}^*(\mathbf{k},t) = -\boldsymbol{\beta}(-\mathbf{k},t) \ . \tag{2.27}$$

Ces expressions prouvent que les variables $\boldsymbol{\beta}$ sont déterminées de façon univoque dès que l'on connaît les $\boldsymbol{\alpha}$, qui sont donc les seuls paramètres dynamiques libres ou encore les seuls degrés de liberté du problème. Nous allons donc en écrire les équations dynamiques.

2.2.2 Dynamique des variables normales

On écrit $\dot{\boldsymbol{\alpha}}$ par simple dérivation et on remplace les dérivées temporelles des champs électriques et magnétiques par les équations de Maxwell correspondantes. On a donc:

$$\begin{aligned} \dot{\boldsymbol{\alpha}} &= -\frac{i}{2\mathcal{N}}(\dot{\boldsymbol{\mathcal{E}}} - c\boldsymbol{\kappa} \times \dot{\boldsymbol{\mathcal{B}}}) \\ &= \frac{ck}{2\mathcal{N}}[c\boldsymbol{\kappa} \times \boldsymbol{\mathcal{B}} + \boldsymbol{\kappa} \times (\boldsymbol{\kappa} \times \boldsymbol{\mathcal{E}})] \ . \end{aligned} \tag{2.28}$$

Le dernier terme dans le crochet de la seconde ligne est simplement égal à $-\boldsymbol{\mathcal{E}}$ puisque le champ est, dans le vide, purement transverse. On a donc finalement:

$$\dot{\boldsymbol{\alpha}} = -ick\boldsymbol{\alpha} = -i\omega\boldsymbol{\alpha} \ , \tag{2.29}$$

en posant, une fois de plus, $\omega = ck$.

Pour comparer cette équation à celle de la dynamique d'un oscillateur, nous considérerons un mouvement unidimensionnel avec la coordonnée x et l'impulsion conjuguée p. Le hamiltonien s'écrit $H = p^2/2m + m\omega x^2/2$. En posant

$$\alpha = x + ip/m\omega \ , \tag{2.30}$$

[1]Sa valeur ne doit être précisée que quand on procède à la quantification du champ électromagnétique. On choisit alors le facteur $\mathcal{N}$, en fonction du volume dans lequel on quantifie le champ, pour que l'énergie d'un photon unique soit précisément égale à $h\nu$.

(on notera la similarité avec les variables normales du champ), on trouve:

$$\dot{\alpha} = -i\omega\alpha \ , \tag{2.31}$$

précisément l'équation décrivant l'évolution de la variable normale du champ.

Chaque vecteur d'onde est donc associé à une variable normale (à deux dimensions puisque transverse) qui a la même dynamique qu'un oscillateur harmonique. Notons que l'on peut se ramener à un problème à une dimension. Pour chaque $\mathbf{k}$, on peut définir deux vecteurs de base orthogonaux à $\mathbf{k}$, $\boldsymbol{\varepsilon}$ et $\boldsymbol{\varepsilon}'$[2]. En écrivant que $\boldsymbol{\alpha} = \alpha_{\boldsymbol{\varepsilon}}\boldsymbol{\varepsilon} + \alpha_{\boldsymbol{\varepsilon}'}\boldsymbol{\varepsilon}'$, on montre que les deux variables normales scalaires $\alpha_{\boldsymbol{\varepsilon}}$ et $\alpha_{\boldsymbol{\varepsilon}'}$ ont la dynamique d'un oscillateur harmonique unidimensionnel. Si on appelle *mode du champ* une valeur de $\mathbf{k}$ et une polarisation, chaque mode du champ électromagnétique est un oscillateur harmonique.

Nous avons franchi là une étape essentielle vers la quantification du champ. On apprend, dans les premiers cours de mécanique quantique, à quantifier l'oscillateur harmonique. Ses niveaux d'énergie sont non dégénérés, égaux à $(n + \frac{1}{2})\hbar\omega$, où n est un nombre entier positif ou nul qui décrit le nombre d'excitations élémentaires ou quanta ou phonons. L'énergie d'un mode quantifié du champ prend donc le même ensemble de valeurs. Le nombre n décrit alors le nombre de quanta d'énergie électromagnétique, le nombre de photons en un mot. Comme toujours en mécanique quantique, les variables classiques sont remplacées par des opérateurs. Les variables $\alpha_{\boldsymbol{\varepsilon}}$ sont remplacées par les opérateurs d'annihilation de photons qui transforment l'état $|n\rangle$ à n photons en l'état $\sqrt{n}|n-1\rangle$ à $n-1$ photons (l'opérateur adjoint $a^\dagger$ étant évidemment l'opérateur de création de photons).

Tant que la quantification se limite à un seul mode (à une seule valeur de $\mathbf{k}$ et de $\boldsymbol{\varepsilon}$), il n'y a aucune difficulté. La plupart des expériences modernes d'optique quantique peuvent se décrire assez précisément en termes de photons et des opérateurs a et $a^\dagger$. En effet, dans la plupart de ces expériences, on ne s'intéresse qu'à quelques modes, définis par des faisceaux laser ou par des cavités résonnantes. La situation se complique un peu quand on veut considérer l'espace libre. Il faut alors sommer tous les modes. L'énergie du fondamental (aucun photon dans aucun mode) ou du vide de rayonnement devient alors infinie. Il s'agit en effet de sommer les énergies de point 0, $1/2\hbar\omega$, d'une infinité de modes. Un zéro d'énergie infini ne serait pas trop grave mais il y a des problèmes plus sérieux. L'effet de ce continuum de modes vides sur un atome, par exemple, est divergent: les niveaux d'énergie sont infiniment modifiés par ce couplage. Il faut donc se débarrasser soigneusement de ces infinis, de ces divergences dues à des sommations sur tous les modes. Les techniques modernes de théorie des champs, de renormalisation, le permettent. En un mot (plus que qualitatif) on peut comprendre qu'une particule chargée isolée du champ n'existe pas. Un électron, par exemple, est "habillé" par les fluctuations de vide. Cet habillage produit une modification infinie des paramètres (masse, charge) de la particule "nue". Mais ce que nous mesurons dans une expérience, la masse ou la charge physique, sont des quantités finies qui tiennent compte de cet habillage. En introduisant dans la théorie des paramètres qui ne sont pas ceux d'une charge isolée mais d'une charge en interaction avec le vide de rayonnement, on peut se débarrasser de ces infinis et obtenir même des prédictions quantitatives d'une précision remarquable.

2.2.3 Décomposition en ondes planes

Ayant explicitement la dynamique des variables normales, nous pouvons, toujours dans l'espace libre, écrire la dynamique du champ dans l'espace réel. On a:

$$\boldsymbol{\alpha} = \alpha_{\boldsymbol{\varepsilon}}(\mathbf{k},0)e^{-i\omega t}\boldsymbol{\varepsilon} + \alpha_{\boldsymbol{\varepsilon}'}(\mathbf{k},0)e^{-i\omega t}\boldsymbol{\varepsilon}' \ . \tag{2.32}$$

[2]Nous prenons ici une base de polarisations linéaires, ce qui est le plus commode à écrire. On pourrait aussi décomposer les variables normales sur une base de polarisations circulaires en prenant une base constituée de superpositions linéaires à coefficient complexes de nos vecteurs.

En portant cette expression dans celle du champ transverse:

$$\boldsymbol{\mathcal{E}} = i\mathcal{N}[\boldsymbol{\alpha}(\mathbf{k},t) - \boldsymbol{\alpha}^*(-\mathbf{k},t)] \tag{2.33}$$

(nous avons utilisé le lien entre $\boldsymbol{\beta}$ et $\boldsymbol{\alpha}$) et en faisant la transformée de Fourier inverse, on trouve finalement:

$$\mathbf{E} = 2\mathrm{Re}\,\frac{i}{(2\pi)^{3/2}}\int d^3\mathbf{k}\,\mathcal{N}(\mathbf{k})\sum_{\boldsymbol{\varepsilon}}\alpha_{\boldsymbol{\varepsilon}}\boldsymbol{\varepsilon}e^{i(\mathbf{k}\cdot\mathbf{r}-\omega t)}\ , \tag{2.34}$$

où la somme sous l'intégrale porte sur deux polarisations transverses arbitraires pour chaque valeur de $\mathbf{k}$. Cette expression est sans surprise. On retrouve, par des voies un peu détournées, le fait que le champ dans l'espace libre peut se décomposer comme une somme d'ondes planes, transverses, de polarisations linéaires.

Appendice 3

Faisceaux gaussiens

Nous chercherons à donner, dans cet appendice, un modèle raisonnable de faisceau laser. Très clairement, l'onde plane ne convient pas pour décrire ces faisceaux. Ils sont bien localisés transversalement. Leur divergence, assez faible, est de l'ordre de la limite imposée par les lois de la diffraction et leur extension transverse. Quand on s'éloigne de l'axe, l'intensité est d'abord stationnaire, puis décroît lentement. Enfin, quand on focalise un tel faisceau au moyen d'une lentille, il existe une taille minimale à la tache focale. Nous allons essayer, à partir de la théorie de la diffraction, de construire un modèle raisonnable pour un tel faisceau dans le premier paragraphe. Nous adopterons une démarche pragmatique mais on pourrait la rendre plus rigoureuse et plus convaincante en montrant qu'on construit ainsi, au moins dans le cadre d'une approximation paraxiale, une solution explicite des équations de Maxwell dans l'espace libre.

Le deuxième paragraphe sera consacré à une revue des propriétés essentielles des modes gaussiens. Nous nous pencherons aussi brièvement sur les modes d'ordre supérieur. Nous discuterons ensuite, dans le paragraphe suivant, des liens entre faisceaux gaussiens et cavités laser. Nous montrerons en particulier qu'une superposition stationnaire de deux faisceaux gaussiens respecte naturellement les conditions aux limites imposées par deux miroirs sphériques dans une configuration Fabry Perot. Nous aurons alors compris pourquoi les lasers produisent effectivement le plus souvent des modes gaussiens. Enfin, nous conclurons ce chapitre par un bref survol de l'optique des modes gaussiens, en nous penchant en particulier sur les transformations de modes au passage par une lentille.

3.1 Construction d'un mode gaussien

Les modes laser sont relativement invariants par translation. Ils divergent légèrement mais la forme de la répartition d'amplitude dans un plan perpendiculaire à l'axe change peu. Or les modifications de cette forme peuvent être décrits par la théorie de la diffraction. L'amplitude en un plan lointain peut se calculer, en première approximation, comme la transformée de Fourier de l'amplitude en un plan proche de l'origine. La seule fonction dont la forme reste invariante par transformation de Fourier est la gaussienne. Il est donc naturel d'essayer de construire un faisceau dont la répartition d'amplitude dans un plan transverse soit gaussienne.

Nous considérerons donc un faisceau d'axe optique Oz. Un plan transverse passant par l'origine sera repéré par les axes Ox_1 et Oy_1. Nous noterons ϕ l'amplitude du champ ou plutôt d'une des composantes du potentiel. Nous supposerons donc que, dans le plan passant par O, la répartition d'amplitude est de la forme:

$$\phi(x_1, y_1) = \phi_0 e^{-(x_1^2+y_1^2)/w_0^2} \ . \tag{3.1}$$

Dans tous les calculs, nous omettrons toujours les facteurs $\exp -i\omega t$. Il est évident, surtout pour des faisceaux laser, que nous ne considérerons que des champs monochromatiques. La dépendance temporelle est alors simplement en facteur de toutes nos expressions et peut être simplement omise. ϕ_0

est une amplitude que nous pouvons supposer réelle sans restreindre la généralité. Nous appellerons, pour des raisons qui seront évidentes dans le prochain paragraphe, w_0 le "waist" du faisceau. Nous supposerons le waist notablement plus grand que λ, longueur d'onde du faisceau. Si ce n'était pas le cas, aucun calcul analytique ne serait possible à partir de la théorie de la diffraction.

Nous chercherons, à partir de là, à calculer la répartition d'amplitude dans un plan Mxy, perpendiculaire à Oz, M étant situé à la cote z. Nous ne nous cantonnerons pas pour ce calcul au régime de Fraunhofer. Nous laisserons, à titre d'exercice, le lecteur le traiter. Il est évident que l'on trouvera une répartition gaussienne de l'intensité, avec une largeur variant linéairement avec la distance z. Cette approximation ne permet pas de définir correctement le faisceau pour des distances z de l'ordre de w_0. Nous allons voir que le principe de Huygens, dans ce cas très simple, permet de s'affranchir de l'approximation de Fraunhofer et d'effectuer un calcul plus précis, valable dans cette région. Bien sûr, nous retrouverons les résultats de Fraunhofer pour des distances suffisamment grandes.

Nous avons en effet:

$$\phi(x,y) = \frac{1}{i\lambda}\int dx_1 dy_1\ \chi\,\frac{\phi(x_1,y_1)}{|\mathbf{r}-\mathbf{r}_1|}\,e^{ik|\mathbf{r}-\mathbf{r}_1|}\ , \tag{3.2}$$

avec les notations standard du chapitre 2. Les vecteurs $\mathbf{r}_1$ et $\mathbf{r}$ sont ici de coordonnées respectives $(x_1, y_1, 0)$ et (x, y, z). Nous allons maintenant approcher l'intégrande en supposant que l'observation n'est pas à trop courte distance. Comme pour la diffraction de Fraunhofer, nous allons traiter très cavalièrement les termes d'amplitude en confondant $|\mathbf{r}-\mathbf{r}_1|$ avec $|z|$ (approximation qui tiendra si la distance d'observation n'est pas trop petite par rapport au waist). Le faisceau est évidemment symétrique par rapport au plan $z = 0$. Nous supposerons donc dans la suite $z > 0$ et reconstruirons ensuite l'autre partie du faisceau par symétrie. Nous traiterons aussi cavalièrement le facteur χ que nous supposerons égal à 1, ce qui est sans doute assez bien vérifié dans le cadre d'une approximation paraxiale. En revanche, nous développerons la phase à un ordre plus élevé que pour le calcul de Fraunhofer:

$$ik|\mathbf{r}-\mathbf{r}_1| \simeq ik\left(z + \frac{(x-x_1)^2}{2z} + \frac{(y-y_1)^2}{2z}\right)\ . \tag{3.3}$$

En tenant compte de la répartition supposée dans le plan Ox_1y_1, on a:

$$\phi(x,y) = \frac{\phi_0}{i\lambda}\frac{e^{ikz}}{z}\int dx_1 dy_1 \exp\left(-\frac{x_1^2+y_1^2}{w_0^2}\right)\exp ik\left[\frac{(x-x_1)^2}{2z} + \frac{(y-y_1)^2}{2z}\right]\ . \tag{3.4}$$

Cette intégrale est explicitement calculable, puisqu'elle ne fait intervenir que des facteurs gaussiens. Pour faire le calcul, nous allons réarranger les termes en x pour faire apparaître un carré. On se convaincra sans problèmes que:

$$\begin{aligned}
-\frac{x_1^2}{w_0^2} + ik\frac{(x-x_1)^2}{2z} &= \frac{ik}{2z}\left(1+\frac{2iz}{kw_0^2}\right)\left[x_1^2 - \frac{2xx_1}{1+2iz/kw_0^2} + \frac{x^2}{1+2iz/kw_0^2}\right] \\
&= \frac{ik}{2z}\left(1+\frac{2iz}{kw_0^2}\right)\left[x_1 - \frac{x}{1+2iz/kw_0^2}\right]^2 - \frac{ik}{2z}\frac{x^2}{1+2iz/kw_0^2} + \frac{ik}{2z}x^2
\end{aligned} \tag{3.5}$$

Les termes en x^2 sont des constantes pour l'intégration qui nous occupe et peuvent donc être mis de côté. En les regroupant et en se livrant à des opérations identiques sur les z, on peut écrire:

$$\begin{aligned}
\phi(x,y) &= \frac{\phi_0}{i\lambda}\frac{e^{ikz}}{z}\exp\left(-\frac{x^2+y^2}{w_0^2(1+2iz/kw_0^2)}\right)\int dx_1 dy_1 \\
&\quad \exp\left[\frac{ik}{2z}(1+2iz/kw_0^2)\left(x_1 - \frac{x}{1+2iz/kw_0^2}\right)^2\right] \\
&\quad \exp\left[\frac{ik}{2z}(1+2iz/kw_0^2)\left(y_1 - \frac{y}{1+2iz/kw_0^2}\right)^2\right]\ .
\end{aligned} \tag{3.6}$$

En posant $u = x_1 - x/(1 + 2iz/kw_0^2)$, l'intégrale sur x_1 se ramène à:

$$\int_D du \exp\left[\left(\frac{ik}{2z} - \frac{1}{w_0^2}\right) u^2\right] , \tag{3.7}$$

le contour d'intégration dans le plan complexe étant la droite D des points de valeur imaginaire égale à Im $(x/(1 + 2iz/kw_0^2))$. Considérons le contour fermé $\mathcal{C}$ constitué de D, de l'axe réel parcouru dans le sens négatif et de deux portions de cercle de l'infini. La fonction à intégrer n'admettant aucun pôle dans le plan complexe, l'intégrale sur de contour est nulle. La contribution des portions à l'infini étant nulle elle aussi, l'intégrale sur D est égale à l'intégrale de la même fonction sur l'axe réel. Or

$$\int_{-\infty}^{\infty} du \exp(-\alpha u^2) = \sqrt{\pi/\alpha} \tag{3.8}$$

même si α est complexe, à condition que les intégrales soient convergentes. Le produit des intégrales sur x_1 et y_1 est donc simplement égal à $\pi/[1/w_0^2 - ik/2z]$ et on a:

$$\phi = \frac{e^{ikz}}{i\lambda z} \frac{\pi}{1/w_0^2 - ik/2z} \exp\left(-\frac{x^2 + y^2}{w_0^2(1 + 2iz/kw_0^2)}\right) , \tag{3.9}$$

ce qui achève notre calcul du point de vue formel. Avant d'aller plus loin, notons que, comme nous l'espérions, la répartition d'amplitude est essentiellement gaussienne.

Pour alléger les notations, nous poserons:

$$b = \frac{kw_0^2}{2} = \frac{\pi w_0^2}{\lambda} . \tag{3.10}$$

Ce paramètre joue visiblement un rôle central dans nos équations. Nous l'appellerons, pour des raisons qui apparaîtront clairement dans l'avant dernier paragraphe de cet appendice, paramètre confocal. On peut alors définir deux longueurs utiles:

$$w(z) = w_0 \sqrt{1 + \frac{z^2}{b^2}} \tag{3.11}$$

et

$$R(z) = \frac{b^2}{z} + z . \tag{3.12}$$

On définira aussi une phase par

$$\Psi = \arctan\left[\frac{z}{b}\right] . \tag{3.13}$$

Avec ces définitions et un peu d'algèbre, on mettra le champ sous la forme

$$\phi = e^{ikz} \frac{w_0}{w(z)} e^{-i\Psi} e^{-(x^2+y^2)/w^2(z)} e^{ik(x^2+y^2)/2R(z)} \tag{3.14}$$

Avant de discuter plus en détails les propriétés de ce faisceau, dont la structure gaussienne est évidente, il faudrait préciser ce qu'est la validité de nos approximations. A priori, notre calcul n'est correct que pour des distance z grandes par rapport au waist. En fait, l'expression que nous obtenons est meilleure. Elle décrit correctement le faisceau pour toutes les distances. On pourra, pour s'en convaincre, porter directement cette expression dans l'équation de propagation des potentiels. On verra qu'elle est une solution de cette équation à condition de faire une approximation paraxiale (on suppose les coordonnées transverses lentement variables par rapport à la longueur d'onde ou les angles de diffraction faibles ce qui revient bien sûr au même). Nous avons donc, un peu par chance, exhibé une solution approximative des équations de Maxwell qui présente toutes les caractéristiques que nous

souhaitions. Nous aurions pu, de manière plus systématique, partir des équations de propagation et de l'approximation paraxiale et en construire une solution. Nous aurions alors trouvé tout un ensemble de solutions, dont celle que nous venons d'exhiber. Les autres, appelées faisceaux gaussiens d'ordre supérieur, seront brièvement évoquées dans le paragraphe suivant.

Nous n'avons ici calculé, a priori, qu'une des composantes du potentiel. Il faudrait les écrire toutes et dériver correctement le potentiel pour trouver la structure des champs électriques et magnétiques. Si la divergence du faisceau n'est pas trop grande, on peut toutefois simplifier beaucoup le calcul. Toutes les quantités relatives au faisceau varient en effet sur une échelle de longueur beaucoup plus grande que λ. On pourra donc sans doute assimiler localement le faisceau gaussien à une onde plane. Les champs électriques et magnétiques seront perpendiculaires à la direction de propagation locale qui est perpendiculaire aux surfaces d'onde ou surfaces d'égale phase du faisceau. Comme les ouvertures sont faibles, les champs ont pratiquement une orientation constante sur le profil du faisceau. En dépit de la géométrie plus complexe, on peut donc continuer à parler d'une polarisation unique pour notre faisceau.

Notons enfin qu'on peut construire des solutions exactes des équations de Maxwell qui se réduisent dans le cadre paraxial aux modes gaussiens ordinaires mais qui les généralisent pour des divergences arbitraires. On les nomme "complex source points" ou CSP. Elles sont obtenues en effet en calculant formellement le rayonnement émis par un dipôle ponctuel situé en iz le long de l'axe Oz. On obtient alors une solution complètement vectorielle des équations de Maxwell. La forme détaillée est trop complexe pour trouver sa place ici.

3.2 Propriétés essentielles

L'interprétation physique de la formule (3.14) est relativement transparente. Le facteur $\exp(-(x^2 + y^2)/w^2(z))$ décrit une répartition gaussienne de l'intensité avec une largeur $w(z)$. La fonction w, normalisée à w_0, est représentée sur la figure 3.1(a). Bien que nous n'ayons fait formellement le calcul que pour $z > 0$, il est clair qu'il peut être étendu dans la région $z < 0$ et que le faisceau est symétrique par rapport à l'origine. Bien sûr, la largeur vaut exactement w_0 en $z = 0$. Ensuite, le faisceau s'élargit, une conséquence naturelle de la diffraction. A grande distance, la taille du faisceau est pratiquement une fonction linéaire de z, correspondant à une ouverture angulaire constante. Cette ouverture angulaire est égale à $2\lambda/\pi w_0$, très voisine de ce qu'on obtient par des arguments qualitatifs de diffraction à la Fraunhofer. On notera en particulier que l'extension latérale du faisceau est multipliée par $\sqrt{2}$ pour $z = b$. Si on imagine le faisceau à trois dimensions, on comprendra bien pourquoi le paramètre w_0 est appelé waist ("taille" en anglais).

Le facteur $\exp(ik(x^2+y^2)/2R(z))$ dans (3.14) affecte la phase du champ. Il coïncide avec la phase d'une onde sphérique de rayon $R(z)$. La surface équiphase, ou surface d'onde, passant par z est donc une sphère de rayon R (à des petites corrections près dues au terme en Ψ). Le rayon R est représenté, normalisé à b, en fonction de z/b sur la figure 3.1(b). A très grande distance, $R \simeq z$. Le centre de courbure des surfaces d'onde est donc l'origine. On retrouve là aussi une conséquence directe de la formule de Huygens dans le régime du champ lointain. En revanche, à l'origine, R est infini: la surface d'onde est un plan, ce que nous avions effectivement supposé initialement. Le rayon de courbure minimal en module est atteint pour $z = \pm b$ et vaut $2b$. Cela justifie le nom de paramètre confocal choisi pour le paramètre b. Les surfaces d'onde en $z = \pm b$ forment en effet deux sphères de même rayon de courbure, le centre de l'une étant situé au sommet de l'autre. C'est précisément la configuration du Fabry Perot confocal, largement utilisé en optique laser. Nous verrons dans le prochain paragraphe de cet appendice comment relier la géométrie des modes gaussiens à la cavité laser qui les définit.

Le terme $\exp -i\Psi$ est une phase supplémentaire. Le paramètre Ψ est représenté en fonction de z/b sur la figure 3.1(c). Sur une étendue de l'ordre de b, Ψ passe de $-\pi/2$ à $\pi/2$. Ce saut de phase par passage au foyer est en fait bien connu en interférométrie classique. Des expériences telles que celles des demi-lentilles montrent clairement que, pour un faisceau quelconque, la phase subit une rotation

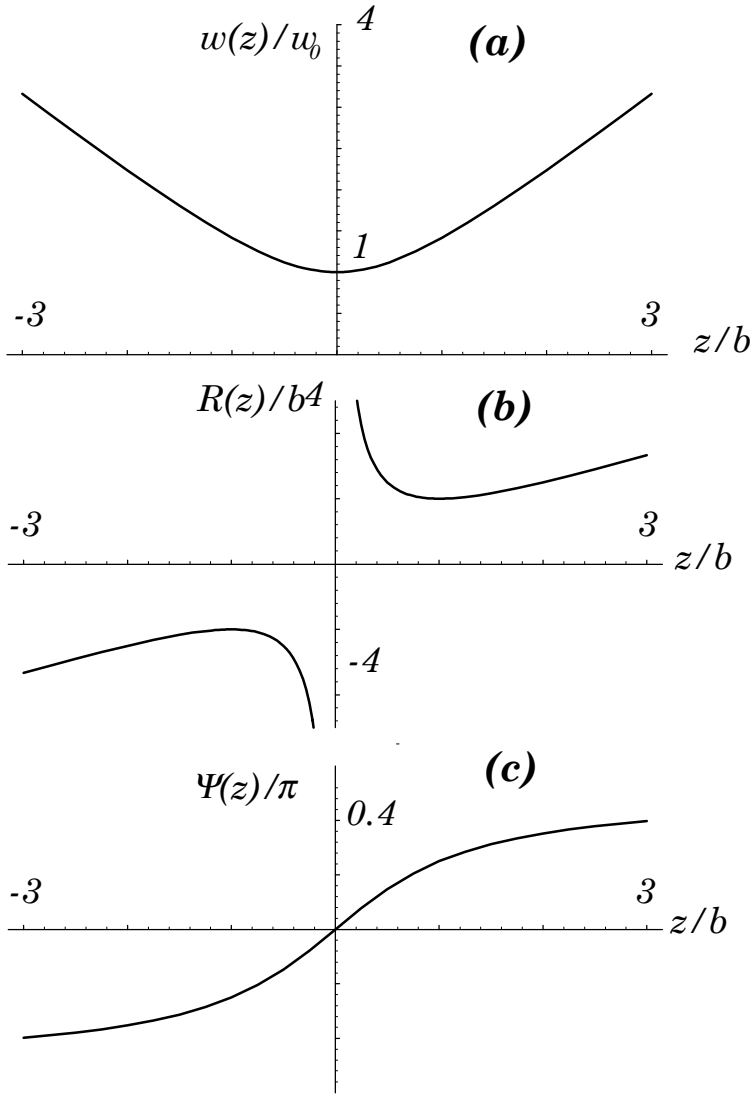

Figure 3.1: (a) Largeur $w(z)$ d'un faisceau gaussien, normalisé à w_0 en fonction de z/b. (b) Rayon de courbure $R(z)$ de la surface d'onde, normalisé à b, en fonction de z/b. (c) Phase Ψ normalisée à π en fonction de z/b.

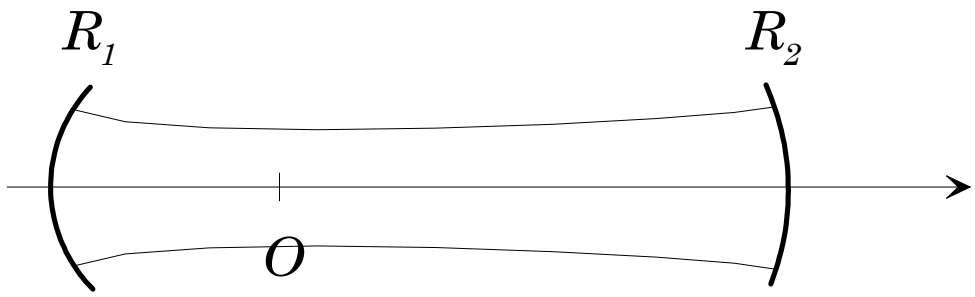

Figure 3.2: Cavité laser Fabry Perot. Elle est constituée de deux miroirs sphériques de rayons de courbure R_1 et R_2 situés en $z_1 < 0$ et $z_2 > 0$. Elle est de longueur L. Elle soutient un mode gaussien dont le waist est situé à l'origine.

de π au passage d'un point focal. On trouvera dans le Born et Wolf des cartes détaillées (et fort complexes) du champ électrique au voisinage d'un foyer pour un faisceau à profil rectangulaire. Ici, la géométrie du faisceau est beaucoup plus simple.

Le dernier terme qui mérite commentaire est le préfacteur $w_0/w(z)$. Il exprime tout simplement la conservation de l'énergie. En prenant le module au carré du "champ", on obtient un terme proportionnel au flux d'énergie. L'intégrale du terme gaussien d'amplitude est alors proportionnel à w. Le préfacteur rétablit la constance de l'énergie transportée le long du faisceau.

Nous n'avons considéré en fait ici que le plus simple des faisceaux gaussiens. Nous aurions pu choisir, dans le plan de référence, une forme plus complexe où la gaussienne serait multipliée par un polynôme de Hermite en fonction de x_1 et de y_1. On retrouve alors un faisceau qui se propage en conservant la même forme de répartition d'amplitude. L'aspect du mode au voisinage de son waist est constitué de taches séparées par des zéros d'intensité correspondant aux zéros des polynômes de Hermite dans les deux directions. Le mode gaussien d'ordre zéro est alors celui qui correspond au polynôme le plus simple, une constante. Les autres faisceaux sont appelés "modes gaussiens d'ordre supérieur". On note souvent TEM_{np} ces faisceaux. TEM désigne un mode transverse électromagnétique. Les champs électriques et magnétiques sont en effet perpendiculaires à la direction de propagation, avec l'approximation paraxiale. n et p désignent les nombres de zéros des polynômes dans les directions x_1 et y_1. Le mode gaussien fondamental est appelé TEM_{00}. C'est en général celui qu'on cherche à réaliser, car il permet, à divergence donnée, d'obtenir la plus petite tache focale. Notons que les différents modes gaussiens forment une base orthogonale de l'ensemble des faisceaux paraxiaux admettant l'axe z comme direction de propagation (le produit scalaire étant l'intégrale de recouvrement d'un des faisceaux avec le complexe conjugué de l'autre). Tout faisceau peut donc a priori se décomposer sur cette base.

3.3 Cavités laser

Un laser est généralement constitué d'une cavité Fabry Perot, plus ou moins remplie par un milieu amplificateur de lumière (voir figure 3.2). Quand le gain sur un aller et retour dans la cavité excède les pertes (à travers les miroirs ou à travers le milieu), l'intensité se renforce dans un aller et retour. On atteint alors le seuil d'oscillation: à partir d'une petite fluctuation de champ, due par exemple à l'émission spontanée par le milieu atomique, l'amplitude croît. Cette divergence est limitée par le phénomène de saturation, qui fait très généralement décroître le gain du milieu amplificateur en fonction de l'intensité présente dans la cavité. On atteint alors un équilibre entre gain et pertes et un régime stationnaire où l'intensité reste constante.

Pour savoir quelle est la géométrie du faisceau produit par un tel laser, on peut chercher à construire un faisceau qui respecte les conditions aux limites imposées par les miroirs. Il faut donc créer une onde stationnaire à partir de deux faisceaux ne différant que par leur direction de propagation dont

les surfaces d'onde soient des sphères, coïncidant avec la surface des miroirs. Les modes gaussiens fondamentaux conviennent a priori. En écrivant la condition d'existence d'une onde stationnaire et l'accord entre le rayon de courbure des surfaces d'onde et celui des miroirs, nous pourrons déterminer complètement la géométrie de ce faisceau à partir de celle de la cavité. C'est à ce calcul que nous allons nous employer.

La géométrie de la cavité est représentée sur la figure 3.2. Nous noterons z_1 la position du miroir 1 par rapport au waist de nos faisceaux gaussiens contrapropageants, z_2 celle du miroir 2. En prenant par convention l'origine de l'axe au waist des faisceaux, nous simplifierons beaucoup les écritures. Toutefois, il ne faut pas perdre de vue que z_1 est alors une inconnue (z_2 étant déterminée à partir de z_1 par la longueur de la cavité). Les rayons de courbure des miroirs seront notés respectivement R_1 et R_2. Dans la configuration représentée sur 3.2, nous conviendrons que les deux rayons de courbure sont positifs. Les résultats que nous allons obtenir sont plus généraux que cette géométrie et admettraient que l'un ou l'autre des rayons de courbure soit négatif.

La première condition à écrire est que l'on puisse former une onde stationnaire entre les deux miroirs. Les deux faisceaux ayant des géométries identiques, il suffit d'écrire les conditions pour le faisceau aller. Si les surfaces d'onde sont confondues avec les miroirs, ce que nous imposerons dans un moment, il suffit d'écrire cette condition de phase aux sommets. La variation de phase sur le faisceau aller, entre z_1 et z_2 doit donc valoir $p\pi$ où p est un entier arbitraire. On doit donc avoir:

$$k(z_2 - z_1) + \Psi(z_2) - \Psi(z_1) = kL + \Psi(z_2) - \Psi(z_1) = p\pi \ . \tag{3.15}$$

L'équation est assez complexe, en raison des termes en Ψ. Elle admet toutefois une seule solution: la longueur de la cavité est fixée de manière univoque quand on fixe p. On retrouve ici tout simplement la condition de résonance de la cavité Fabry Perot. On peut fixer un peu les choses si la distance entre les miroirs et entre miroirs et waist est grande devant le paramètre confocal du faisceau (ce qui n'est guère le cas, comme nous le verrons, pour les cavités standard). On a en effet $\Psi(z_2) - \Psi(z_1) = \pi$ et la condition de résonance s'écrit simplement:

$$L = (p + \frac{1}{2})\frac{\lambda}{2} \ , \tag{3.16}$$

la condition obtenue naïvement en négligeant les termes en Ψ étant $L = p\lambda/2$.

L'accord en rayon de courbure entre les surfaces d'onde et les miroirs s'écrit donc simplement:

$$R(z_1) = -R_1 \tag{3.17}$$

$$R(z_2) = R_2 \ , \tag{3.18}$$

le signe étant imposé par nos conventions de prendre positifs les rayons de courbure des miroirs. Si on y ajoute le fait que la longueur de la cavité est L, nous avons à résoudre le système:

$$z_1 + \frac{b^2}{z_1} = -R_1 \tag{3.19}$$

$$z_2 + \frac{b^2}{z_2} = R_2 \tag{3.20}$$

$$z_2 - z_1 = L \ . \tag{3.21}$$

Les inconnues sont le waist w_0 du faisceau et sa position par rapport aux miroirs. Comme nous avons choisi le waist comme origine, nous pourrons prendre comme inconnues z_1 et z_2 (la troisième équation donnant trivialement l'un en fonction de l'autre). En soustrayant les deux premières équations, en faisant systématiquement intervenir la distance $z_2 - z_1$ c'est à dire L, on écrira sans peine:

$$b^2 = z_1 z_2 \left(1 - \frac{R_1}{L} - \frac{R_2}{L}\right) \ . \tag{3.22}$$

Introduisons alors deux nouveaux paramètres donnés par:

$$g_1 = 1 - \frac{L}{R_1} \tag{3.23}$$

$$g_2 = 1 - \frac{L}{R_2} \ . \tag{3.24}$$

On a $R_1/L = 1/(1-g_1)$ et:

$$b^2 = \frac{z_1 z_2 (g_1 g_2 - 1)}{(1-g_1)(1-g_2)} \tag{3.25}$$

En multipliant la première équation de notre système par z_1, la deuxième par z_2 et en en faisant enfin la différence, nous avons aussi:

$$z_2^2 - z_1^2 = L(z_2 + z_1) = R_2 z_2 + R_1 z_1 \ , \tag{3.26}$$

et

$$(L - R_2) z_2 + (L - R_1) z_1 = 0 \ . \tag{3.27}$$

En divisant par $R_1 R_2$, on en tire sans difficultés:

$$\frac{g_2}{R_1} z_2 + \frac{g_1}{R_2} z_1 = 0 \ . \tag{3.28}$$

En exprimant enfin z_1 en fonction de L et z_2, on obtient

$$z_2 = L \frac{g_1 R_1}{g_2 R_2 + g_1 R_1} \tag{3.29}$$

ou encore

$$z_2 = L \frac{g_1 (1 - g_2)}{g_1 + g_2 - 2 g_1 g_2} \ . \tag{3.30}$$

On en déduit:

$$z_1 = -L \frac{g_2 (1 - g_1)}{g_1 + g_2 - 2 g_1 g_2} \tag{3.31}$$

et, par simple substitution:

$$b^2 = L^2 \frac{g_1 g_2 (1 - g_1 g_2)}{(g_1 + g_2 - 2 g_1 g_2)^2} \ , \tag{3.32}$$

ce qui achève la résolution du système.

Nous avons donc ainsi exprimé la position et la valeur du waist du faisceau en fonction des paramètre géométriques de la cavité. Nous n'aurons toutefois une solution acceptable que si $b^2 > 0$, c'est à dire si:

$$0 < g_1 g_2 < 1 \ . \tag{3.33}$$

Le point représentatif de la cavité dans un plan g_1, g_2 doit donc être dans le premier ou le troisième quadrant, entre les axes et les deux branches de l'hyperbole $g_1 g_2 = 1$. Cette condition est une condition de stabilité pour la cavité laser. On peut en effet, d'un point de vue purement optique, remplacer les trajets aller et retour entre les miroirs de la cavité par une propagation dans une série de lentilles, convergentes ou divergentes selon la courbure des miroirs, espacées de L. Traiter ce petit problème en optique géométrique ne pose aucun problème si on écrit la matrice de transfert correspondant à un passage par deux lentilles consécutives ou un aller et retour dans la cavité. La propagation sur un grand nombre de passages ou d'allers et retour sera décrite par cette matrice élevée à une grande puissance. En général, la position ou l'angle du faisceau avec l'axe, qui sont les paramètres naturels de l'optique matricielle, divergent dans cette opération. C'est seulement, comme on pourra le vérifier, si la condition (3.33) est remplie que le rayon reste à distance finie de l'axe. Cette condition est donc

une pure condition d'optique géométrique qui exprime que les deux miroirs confinent effectivement la lumière au voisinage de l'axe.

On prend donc bien garde, sauf dans quelques cas particuliers (milieux de très grand gain) de réaliser des cavités laser stables. Dans le plan g_1, g_2, nous pouvons examiner quelques points particuliers, à la frontière du domaine de stabilité

- $g_1 = g_2 = 0$. On a alors $R_1 = R_2 = L$. La cavité est une cavité Fabry Perot confocale, le centre de courbure d'un miroir étant confondu avec le sommet de l'autre. C'est, de loin, la configuration de cavité la plus utilisée en technologie laser. Elle présente, entre autres, l'avantage que les modes gaussiens d'ordre supérieur se trouvent dégénérés en fréquence avec le mode fondamental. On peut voir facilement que $b = L/2$, ce dont nous pouvions nous douter, et que $w_0 = \sqrt{L\lambda/2\pi}$. Comme cette cavité est à la limite du domaine de stabilité, on rapproche en général légèrement les deux miroirs.

- $g_1 = g_2 = 1$ Les deux miroirs sont plans. C'est aussi une cavité à la limite de la stabilité, puisqu'on trouve évidemment un waist infini.

- $g_1 = g_2 = -1$ On trouve alors $R_1 = R_2 = L/2$. La cavité est une cavité concentrique. En fait les deux miroirs forment deux calottes d'une même sphère. On trouve dans ce cas, par un passage à la limite sans difficultés, que le waist devrait être nul. Bien sûr, cette solution n'a aucun sens physique, les conditions de l'approximation paraxiale étant violées depuis longtemps.

3.4 Optique gaussienne

On utilise bien sûr les faisceaux gaussiens dans des système optiques composés entre autres de lentilles minces. Il est donc important de savoir comment se transforme le faisceau au passage dans ce composant essentiel de l'optique. Nous allons voir que, si les résultats sont le plus souvent très voisins de ceux de l'optique géométrique standard, ils en sont parfois aussi éloignés de manière non intuitive.

Pour décrire une lentille mince, nous allons examiner son influence sur la phase d'une onde incidente. L'axe de la lentille sera bien sûr confondu avec Oz. Comme la lentille est très mince, la déviation des rayons lumineux est négligeable et la répartition transverse d'amplitude n'est pas modifiée. Nous avons déjà rencontré ces arguments, dans le chapitre 2, quand nous avons introduit la notion d'objet de phase pour la fonction de transmission d'un diaphragme. Considérons, pour fixer les idées, une lentille plan-convexe d'indice n. La face d'entrée est plane, la face de sortie est une sphère de rayon R. L'épaisseur de verre traversée par le faisceau à la distance $x^2 + y^2$ de l'axe est:

$$\ell = \ell_0 - \frac{x^2 + y^2}{2R} \ , \tag{3.34}$$

où ℓ_0 est l'épaisseur maximale (nous avons bien sûr, de manière consistante avec l'approximation paraxiale, remplacé la sphère par une parabole). Nous pouvons omettre ici les termes en ℓ_0, qui décriraient le déphasage d'une lame mince d'épaisseur égale à celle de la lentille. Ils provoquent seulement un déphasage global de l'onde incidente. L'amplitude de l'onde en sortie est donc multipliée par le facteur de phase:

$$\exp\left(-ik(n-1)\frac{x^2+y^2}{2R}\right) = \exp\left(-ik\frac{x^2+y^2}{2f}\right) \ . \tag{3.35}$$

Nous avons fait intervenir $f = R/(n-1)$. Pour interpréter ce résultat, on pourra s'apercevoir qu'une onde incidente plane émerge avec la phase d'une onde sphérique convergente de rayon de courbure f. f est donc la distance focale de la lentille, un résultat élémentaire d'optique géométrique. Le résultat

que nous venons d'obtenir avec une lentille convergente peut se généraliser à toutes les lentilles minces. Il suffit, avec nos notations, de compter f négativement quand la lentille est divergente.

Considérons maintenant un faisceau gaussien incident. Il est décrit, dans le plan de la lentille, par $w(z)$ et $R(z)$ ainsi que par la phase $\Psi(z)$. La répartition transverse d'intensité n'étant pas modifiée, l'amplitude reste une fonction gaussienne de la distance à l'axe, de même largeur. En revanche, la phase due à la lentille s'ajoute à celle du faisceau. L'onde résultante est encore une onde sphérique, avec une courbure R' telle que:

$$\frac{1}{R'} = \frac{1}{R} - \frac{1}{f} \ . \tag{3.36}$$

Il est donc évident que le faisceau émergent est lui aussi gaussien, décrit par des paramètres $w'(z')$ et $R'(z')$, z' étant la distance du waist (encore inconnu) de ce faisceau à la lentille. Notons à ce point que la phase Ψ d'un faisceau gaussien peut se calculer facilement à partir de w et R. Toute l'influence de la phase de la lentille ayant été inclus dans le terme en R', la phase Ψ' est évidemment égale à Ψ. En général, Ψ' ne coïncide pas avec la phase du nouveau faisceau calculée à partir de w' et R'. Il peut donc exister en plus du changement de géométrie un déphasage global. Les nouveaux paramètres du faisceau peuvent se calculer en résolvant l'équation de conjugaison (3.36) et $w' = w$. On en déduit la cote z' et le nouveau paramètre confocal b' qui fixent la position et l'étendue du nouveau waist.

Ce calcul est un peu complexe, faisant intervenir des équations du quatrième ordre. Nous n'examinerons que quelques cas simples. Le faisceau incident a son waist à l'origine que nous supposerons éloignée de la lentille. Le rayon de courbure R de ce faisceau est donc essentiellement égal à la cote z de la lentille. Considérons d'abord le cas où R' s'annule. Il est évident que le waist émergent est situé sur le plan de la lentille. Cela se produit, d'après l'équation de conjugaison, si $R = f$, c'est à dire si le waist du faisceau incident est situé au foyer de la lentille. En optique géométrique, l'image de ce plan est à l'infini. Nous voyons bien ici que l'optique gaussienne n'est pas toujours intuitive et qu'en particulier le waist du faisceau image n'est pas toujours confondu avec l'image géométrique du plan du waist d'entrée. En revanche, en supposant que le waist du faisceau émergent est, lui aussi, situé loin de la lentille, le rayon de courbure R' donne directement sa position qui coïncide avec l'image géométrique, comme le montre la relation de conjugaison. On peut montrer aussi que, dans ce cas, les deux waists sont dans un rapport égal au grandissement géométrique de la lentille.

Appendice 4

Approximation eikonale

Nous avons vu, à la fin du chapitre trois, comment on pouvait passer qualitativement de l'électromagnétisme à la notion de rayon lumineux en utilisant les concepts de la diffraction. Nous avons vu aussi quels liens existaient alors entre le principe de Fermat et cette approche. Nous allons, dans cet appendice, rendre beaucoup plus rigoureuse cette approche. Nous partirons des équations de Maxwell et nous ferons un changement de variables, tout à fait rigoureux, pour distinguer deux échelles de longueur dans la variation des champs. Nous aurons d'une part la longueur d'onde dont l'échelle décrit les variations de phase du champ dans la direction de la propagation. Nous aurons aussi la variation spatiale "transverse" des champs due au fait que nous ne propageons pas en général des ondes planes. Nous ferons alors, dans le deuxième paragraphe, l'approximation de l'optique géométrique qui revient à négliger l'échelle spatiale de la première variation devant la seconde. En supposant ainsi que les "enveloppes" des champs varient lentement devant la longueur d'onde, nous retrouverons la notion de rayon lumineux, clairement définie en examinant le transport d'énergie. Enfin, dans le troisième paragraphe, nous écrirons l'équation décrivant la trajectoire des rayons lumineux. Nous montrerons aussi que cette équation est celle que l'on obtient directement à partir du principe de Fermat. Nous nous livrerons pour cela à un calcul variationnel très inspiré de ceux de la première partie de ce cours.

Notons que la technique que nous allons employer ici pour passer à l'optique géométrique s'utilise aussi dans des domaines variés. En mécanique quantique, par exemple, de nombreuses méthodes d'approximations semi–classiques sont fondées sur ce genre d'approche. Si les calculs ne sont pas directement transposables (ils sont relativement plus complexes ici en raison de la nature vectorielle du champ), leur esprit reste similaire.

Dans tout cet appendice, nous ne considérerons que la propagation dans un milieu diélectrique, sans propriétés magnétiques (les milieux magnétiques étant en général opaques, l'optique y est sans grand intérêt). Nous ne considérerons aussi que des champs monochromatiques, oscillant à la fréquence ω. Nous noterons ϵ_r la permittivité diélectrique du milieu et $n = \sqrt{\epsilon_r}$ son indice de réfraction à cette même fréquence. Nous supposerons, pour faciliter les calculs, ces quantités réelles et scalaires. Nous ne traiterons donc que de milieux sans absorption ni biréfringence (des généralisations relativement immédiates permettent de tenir compte de ces effets). En revanche, nous ne ferons aucune hypothèse sur l'homogénéité du milieu et permettrons à ces quantités de varier avec la position. Les différentes quantités décrites ici sont définies et commentées dans la cinquième partie de ce cours sur la propagation dans les milieux matériels. Enfin, nous ne traiterons que la propagation libre indépendamment des sources. C'est bien en général le problème posé par une expérience d'optique.

4.1 Eikonale

Nous partons donc des équations de Maxwell pour des champs monochromatiques. Elles s'écrivent, en notant $\mathbf{E}_0(\mathbf{r})$ et $\mathbf{B}_0(\mathbf{r})$ les amplitudes complexes des champs électrique et magnétique en un point

(nous nous débarrassons ainsi de toute la dépendance temporelle qui est en facteur dans tout ce qui va suivre):

$$\begin{aligned}
\boldsymbol{\nabla} \times \mathbf{E}_0 &= ick_0\mathbf{B}_0 && (4.1)\\
\boldsymbol{\nabla} \cdot \epsilon_r\mathbf{E}_0 &= 0 && (4.2)\\
\boldsymbol{\nabla} \cdot \mathbf{B}_0 &= 0 && (4.3)\\
\boldsymbol{\nabla} \times \mathbf{B}_0 &= -i\frac{k_0}{c}\mathbf{E}_0 \ , && (4.4)
\end{aligned}$$

en posant $k_0 = \omega/c$.

Si nous imaginons ce que doit être un rayon lumineux, la phase des champs dans la direction de propagation doit varier rapidement (sur une échelle $1/k_0$) alors que, perpendiculairement à cette direction, l'amplitude varie lentement. Nous allons donc changer de variables pour faire apparaître d'une part une variation rapide de la phase avec une échelle k_0 et d'autre part une variation plus lente de l'amplitude. En un mot, nous faisons apparaître les enveloppes lentement variables des champs d'une part et la modulation rapide due à leur caractère oscillant d'autre part. Ce genre d'approximations d'enveloppes lentement variables est assez couramment utilisée. Nous poserons donc:

$$\begin{aligned}
\mathbf{E}_0 &= \mathbf{e}(\mathbf{r})e^{ik_0\mathcal{S}(\mathbf{r})} && (4.5)\\
\mathbf{B}_0 &= \mathbf{b}(\mathbf{r})e^{ik_0\mathcal{S}(\mathbf{r})} && (4.6)\\
& && (4.7)
\end{aligned}$$

où $\mathcal{S}$, appelée eikonale, est une fonction réelle de $\mathbf{r}$. Ce changement de variable n'est bien sûr pas univoque. Il n'a pas vocation à l'être, puisque nous n'étudions que la propagation dans l'espace libre, indépendamment des sources. Le champ ne serait complètement défini que si nous pouvions spécifier aussi les sources. Pour le débat qui nous occupe maintenant, la définition de la notion de rayon lumineux dans la propagation libre, cette question est sans intérêt. Nous pouvons néanmoins constater que les champs des sources à grande distance, qui seront détaillés dans la partie suivante, peuvent tous se mettre sous la forme précédente avec un choix très naturel pour l'eikonale.

Porter ce changement de variables dans les équations de Maxwell ne pose aucun problème technique. Pour la première on a:

$$\begin{aligned}
\boldsymbol{\nabla} \times \mathbf{e}e^{ik_0\mathcal{S}} &= \boldsymbol{\nabla}e^{ik_0\mathcal{S}} \times \mathbf{e} + e^{ik_0\mathcal{S}}\boldsymbol{\nabla} \times \mathbf{e}\\
&= ik_0(\boldsymbol{\nabla}\mathcal{S})e^{ik_0\mathcal{S}} \times \mathbf{e} + e^{ik_0\mathcal{S}}\boldsymbol{\nabla} \times \mathbf{e} \ . && (4.8)
\end{aligned}$$

La première équation de Maxwell s'écrit donc:

$$\boldsymbol{\nabla}\mathcal{S} \times \mathbf{e} + \frac{\boldsymbol{\nabla} \times \mathbf{e}}{ik_0} = c\mathbf{b} \ . \tag{4.9}$$

Pour la deuxième:

$$\boldsymbol{\nabla} \cdot \epsilon_r\mathbf{e}e^{ik_0\mathcal{S}} = \boldsymbol{\nabla}\epsilon_r \cdot \mathbf{e}e^{ik_0\mathcal{S}} + \epsilon_r\mathbf{e} \cdot ik_0(\boldsymbol{\nabla}\mathcal{S})e^{ik_0\mathcal{S}} + \epsilon_r e^{ik_0\mathcal{S}}\boldsymbol{\nabla} \cdot \mathbf{e} \tag{4.10}$$

et donc

$$\mathbf{e} \cdot \boldsymbol{\nabla}\mathcal{S} = -\frac{1}{ik_0}\left[(\boldsymbol{\nabla} \ln \epsilon_r) \cdot \mathbf{e} + \boldsymbol{\nabla} \cdot \mathbf{e}\right] \tag{4.11}$$

Par des calculs similaires, on montre que les deux dernières équations de Maxwell se mettent sous la forme:

$$\begin{aligned}
\boldsymbol{\nabla}\mathcal{S} \cdot \mathbf{b} &= -\frac{1}{ik_0}\boldsymbol{\nabla} \cdot \mathbf{b} && (4.12)\\
\boldsymbol{\nabla}\mathcal{S} \times \mathbf{b} + \frac{\epsilon_r}{c}\mathbf{e} &= -\frac{1}{ik_0}\boldsymbol{\nabla} \times \mathbf{b} \ . && (4.13)
\end{aligned}$$

Insistons sur le fait que tous ces calculs sont exacts. Bien entendu, les quatre équations de Maxwell ne permettent pas de déterminer les 7 fonctions inconnues de nos nouvelles variables. On retrouve le fait que ce changement de fonctions est ambigu.

4.2 Approximation de l'optique géométrique

Nous allons maintenant faire réellement l'approximation de l'optique géométrique. Nous allons supposer que les enveloppes lentement variables et l'eikonale varient toutes sur une échelle de longueur très supérieures à la longueur d'onde. Dans les équations de Maxwell, coexistent des dérivées de ces quantités par rapport à la position et des termes faisant intervenir $1/k_0$. Avec notre approximation, ces derniers termes sont négligeables par rapport aux premiers (nous laissons la justification mathématique au lecteur!). L'approximation de l'optique géométrique consiste à négliger ces termes et à écrire les équations de Maxwell pour les enveloppes lentement variables sous la forme:

$$\boldsymbol{\nabla}\mathcal{S} \times \mathbf{b} + \frac{\epsilon_r}{c}\mathbf{e} = 0 \tag{4.14}$$

$$\boldsymbol{\nabla}\mathcal{S} \times \mathbf{e} - c\mathbf{b} = 0 \tag{4.15}$$

$$\mathbf{e} \cdot \boldsymbol{\nabla}\mathcal{S} = 0 \tag{4.16}$$

$$\mathbf{b} \cdot \boldsymbol{\nabla}\mathcal{S} = 0 \ . \tag{4.17}$$

Notons tout de suite que les champs électrique et magnétique sont perpendiculaires au gradient de l'eikonale. La phase de l'onde étant essentiellement déterminée par le terme en $k_0\mathcal{S}$, les surfaces d'onde, ou surfaces équiphases, sont perpendiculaires au gradient de $\mathcal{S}$. Le gradient de l'eikonale pointe donc dans la direction locale de propagation. Les champs électrique et magnétique sont donc perpendiculaires à cette direction, contenus dans la surface d'onde. L'énergie électromagnétique se propage donc dans la direction du gradient de l'eikonale, perpendiculairement aux surfaces d'onde. Nous développerons plus avant ces arguments énergétiques dans un moment. Notons aussi que la deuxième équation montre alors que $\mathbf{e}$, $\mathbf{b}$ et $\boldsymbol{\nabla}\mathcal{S}$ forment un trièdre direct. Notre onde a donc localement la structure d'une onde plane de vecteur d'onde effectif $k_0\boldsymbol{\nabla}\mathcal{S}$. Il est évident qu'une approximation de champs lentement variables doit, comme ce sera le cas pour le rayonnement des sources à grande distance, conduire à une structure locale d'onde plane.

4.2.1 Equation de l'eikonale

Les deux premières équations de Maxwell ainsi écrites doivent admettre une solution pour $\mathbf{e}$ et $\mathbf{b}$. Si on se donne $\boldsymbol{\nabla}\mathcal{S}$ en un point, ces deux équations forment, en termes des composantes cartésiennes des champs, un système linéaire homogène 6×6. En général, il n'admet que la solution triviale identiquement nulle. Pour que $\mathbf{e}$ et $\mathbf{b}$ existent dans le cadre de cette approximation, il faut qu'en chaque point le gradient de l'eikonale prenne la valeur particulière qui annule le déterminant de ce système. Plutôt que d'écrire ainsi cette condition d'existence, nous allons procéder par substitution. De la seconde équation, on tire:

$$\mathbf{b} = \boldsymbol{\nabla}\mathcal{S} \times \mathbf{e}/c \ , \tag{4.18}$$

qu'on porte dans la première. On obtient alors:

$$\boldsymbol{\nabla}\mathcal{S} \times (\boldsymbol{\nabla}\mathcal{S} \times \mathbf{e}) + \epsilon_r\mathbf{e}/c = 0 \ . \tag{4.19}$$

En développant le double produit vectoriel et en utilisant le fait que $\mathbf{e}$ est perpendiculaire au gradient de l'eikonale et en introduisant finalement l'indice de réfraction, on met la condition d'existence des champs sous la forme:

$$(\boldsymbol{\nabla}\mathcal{S})^2 = n^2 \tag{4.20}$$

En tous points, l'eikonale doit vérifier cette relation pour que les champs existent. Une partie de l'arbitraire sur sa définition est donc levée.

4.2.2 Notion de rayon lumineux

Pour préciser la notion de rayon lumineux, nous allons nous pencher sur les propriétés énergétiques des champs ainsi définis. Ecrivons d'abord le vecteur de Poynting, en valeur moyenne temporelle:

$$\overline{\mathbf{\Pi}} = \frac{1}{2}\mathrm{Re}\ \frac{\mathbf{e}\times\mathbf{b}^*}{\mu_0} \tag{4.21}$$

où l'étoile désigne la conjugaison complexe. Notons que l'eikonale, pur facteur de phase, s'élimine des quantités énergétiques en valeur moyenne. En exprimant alors le champ magnétique à partir de la deuxième équation de propagation, en développant le double produit vectoriel et en exploitant la géométrie, on a:

$$\overline{\mathbf{\Pi}} = \frac{\epsilon_0 c}{2}\mathrm{Re}\ (\mathbf{e}\cdot\mathbf{e}^*)\boldsymbol{\nabla}\mathcal{S}\ . \tag{4.22}$$

Pour éclairer cette équation, nous allons écrire la densité moyenne d'énergie électromagnétique, $\overline{u}$. On a (voir cinquième partie):

$$\overline{u} = \frac{\epsilon_0\epsilon_r}{4}\mathrm{Re}\ \mathbf{e}\cdot\mathbf{e}^* + \frac{1}{4\mu_0}\mathrm{Re}\ \mathbf{b}\cdot\mathbf{b}^*\ . \tag{4.23}$$

On peut montrer rapidement que les contributions électriques et magnétiques à l'énergie sont égales (notons que c'est toujours le cas pour une onde plane qui est la structure locale de notre champ). En remplaçant, dans l'énergie électrique, $\mathbf{e}^*$ par son expression en fonction de $\mathbf{b}^*$ déduite de (4.14), $\mathbf{e}^* = (c\boldsymbol{\nabla}\mathcal{S}\times\mathbf{b}^*)/\epsilon_r$ (on notera que par définition $\boldsymbol{\nabla}\mathcal{S}$ est réel), on obtient un produit mixte:

$$\frac{\epsilon_0\epsilon_r}{4}\mathrm{Re}\ \mathbf{e}\cdot\mathbf{e}^* = \frac{\epsilon_0 c}{4}[\mathbf{e},\mathbf{b}^*,\boldsymbol{\nabla}\mathcal{S}]\ . \tag{4.24}$$

En remplaçant $\mathbf{b}^*$ par son expression en fonction de $\mathbf{e}^*$ déduite de (4.15) dans l'énergie magnétique, en remarquant que $\mu_0 c = 1/\epsilon_0 c$, on constate sans peine que la contribution magnétique a une expression identique. Finalement, la valeur moyenne de l'énergie totale s'écrit:

$$\overline{u} = \frac{\epsilon_0 n^2}{2}\mathrm{Re}\ \mathbf{e}\cdot\mathbf{e}^*\ . \tag{4.25}$$

En rapprochant cette expression de celle du vecteur de Poynting, on constate que celui-ci est proportionnel à la densité locale d'énergie. Pour alléger les notations, nous appellerons $\mathbf{s}$ le vecteur unitaire du gradient de l'eikonale:

$$\mathbf{s} = \frac{\boldsymbol{\nabla}\mathcal{S}}{|\boldsymbol{\nabla}\mathcal{S}|} = \frac{\boldsymbol{\nabla}\mathcal{S}}{n} \tag{4.26}$$

(on notera ici l'importance de l'hypothèse n réel). On a alors:

$$\overline{\mathbf{\Pi}} = \frac{c}{n}\overline{u}\mathbf{s}\ . \tag{4.27}$$

On a donc bien une propagation de l'énergie dans la direction perpendiculaire aux surfaces d'onde, avec une vitesse c/n puisque le rapport entre le flux d'énergie et sa densité moyenne n'est autre que la vitesse de propagation (une vitesse de groupe en l'occurrence). Nous pouvons donc maintenant assimiler les lignes du champ $\boldsymbol{\nabla}\mathcal{S}$ aux rayons lumineux de l'optique géométrique. A condition de ne pas les limiter à une échelle de l'ordre de λ — et de sortir donc de nos approximations d'enveloppes lentement variables — les lignes de $\boldsymbol{\nabla}\mathcal{S}$ décrivent effectivement les lignes géométriques de propagation de l'énergie. Comme ces lignes sont perpendiculaires aux surfaces d'onde, on obtient ainsi directement le théorème de Malus, qui indique que les rayons lumineux sont perpendiculaires aux surfaces d'onde[1]. Il nous reste maintenant à établir la trajectoire de ces rayons.

[1]Notons que le théorème de Malus ne tient que dans les milieux diélectriques isotropes (dont la permittivité diélectrique est un scalaire). Dans les milieux biréfringents, on a encore localement une structure d'onde plane. Cette fois, $\mathbf{D}$,$\mathbf{B}$ et la direction de propagation de la phase (le vecteur d'onde local) forment un trièdre direct. Mais le vecteur de Poynting est encore déterminé par $\mathbf{E}$. Comme la relation entre $\mathbf{D}$ et $\mathbf{E}$ est tensorielle dans ces milieux, les deux vecteurs ne sont pas nécessairement alignés. Le vecteur de Poynting, qui définit la direction de propagation de l'énergie et donc la direction des rayons lumineux, n'est pas nécessairement aligné avec la direction de propagation de la phase, normale aux surfaces d'onde. Le théorème de Malus est donc en général inapplicable dans ces milieux.

4.3 Equation des rayons lumineux

4.3.1 Rayons et eikonale

Les rayons lumineux sont donc les lignes du champ de gradient de $\mathcal{S}$. Il nous reste à établir la forme de ces lignes. Nous chercherons donc le rayon lumineux passant par un point arbitraire. Le rayon peut être défini par la forme intrinsèque $\mathbf{r}(s)$, position du point en fonction de l'abscisse curviligne s comptée sur le rayon à partir du point initial. Le rayon est effectivement une ligne de $\boldsymbol{\nabla}\mathcal{S}$ si, en chaque point, le vecteur tangent à la trajectoire est confondu avec le vecteur unitaire $\mathbf{s}$ du gradient (on prendra garde à ne pas confondre ce vecteur $\mathbf{s}$ avec l'abscisse curviligne s le long du rayon en dépit d'une fâcheuse homonymie). On doit donc avoir:

$$\mathbf{s} = \frac{d\mathbf{r}}{ds} \ , \tag{4.28}$$

ou encore, en multipliant les deux membres par n, longueur du gradient de l'eikonale:

$$n\frac{d\mathbf{r}}{ds} = \boldsymbol{\nabla}\mathcal{S} \ . \tag{4.29}$$

Essayons d'écrire une équation intrinsèque pour $\mathbf{r}$ en éliminant l'eikonale. Pour cela, afin de manipuler commodément les indices, nous quittons pour un moment les notations standard de l'analyse vectorielle et adoptons les conventions d'Einstein de sommation. Nous écrirons les indices indifféremment en position contravariante ou covariante puisque nous faisons un calcul seulement avec des composantes purement spatiales. Ecrivons donc la quantité:

$$\frac{d}{ds}n\frac{d\mathbf{r}}{ds} \ , \tag{4.30}$$

en distinguant les composantes et en utilisant l'équation (4.29):

$$\frac{d}{ds}n\frac{d\mathbf{r}_i}{ds} = \frac{d}{ds}\partial_i\mathcal{S} \ . \tag{4.31}$$

Mais:

$$\frac{d}{ds}\partial_i\mathcal{S} = \partial_j[\partial_i\mathcal{S}]\frac{d\mathbf{r}^j}{ds} \ . \tag{4.32}$$

En remplaçant dans cette expression $d\mathbf{r}^j/ds$ par $\partial^j\mathcal{S}/n$, on a:

$$\begin{aligned} \frac{d}{ds}n\frac{d\mathbf{r}_i}{ds} &= \frac{1}{n}\partial_j[\partial_i\mathcal{S}]\partial^j\mathcal{S} \\ &= \frac{1}{2n}\partial_i[\partial_j\mathcal{S}\partial^j\mathcal{S}] \\ &= \frac{1}{2n}\partial_i[n^2] \ , \end{aligned} \tag{4.33}$$

où nous avons utilisé l'équation fondamentale de l'eikonale dans le second membre. Finalement, l'équation intrinsèque du rayon lumineux s'écrit donc, en revenant aux notations standard:

$$\frac{d}{ds}n\frac{d\mathbf{r}}{ds} = \boldsymbol{\nabla}n \ . \tag{4.34}$$

Elle permet de déterminer complètement la trajectoire du rayon si on connaît la répartition spatiale d'indice.

4.3.2 Applications

Pour regarder fonctionner cette équation, nous allons examiner rapidement quelques répartitions d'indice assez couramment rencontrées. Notons tout d'abord que notre équation ne s'applique pas a priori au cas d'un indice discontinu, tel que celui au voisinage d'un dioptre. On peut toutefois, en prenant la limite d'une variation continue de plus en plus rapide, en déduire les lois de la réfraction.

Commençons par un milieu homogène, d'indice constant. Le gradient et nul et l'indice sort du premier membre, mettant l'équation du rayon lumineux sous la forme:

$$\frac{d^2\mathbf{r}}{ds^2} = 0 \ , \tag{4.35}$$

dont la solution est évidement une droite. Nous retrouvons donc la propagation rectiligne de la lumière dans un tel milieu.

Considérons maintenant un milieu dont l'indice ne dépend que de la coordonnée z, avec une dépendance linéaire :$n = n_0(1+az)$. C'est par exemple le cas de l'air quand il y a formation de mirage. L'air est beaucoup plus chaud au contact du sol, exposé au soleil, qu'en altitude. La densité, et donc l'indice, sont donc des fonctions croissantes de l'altitude z. Pour fixer les idées, nous considérerons un rayon se propageant de manière pratiquement horizontale dans un plan xOz. Nous pouvons alors, en première approximation, confondre l'abscisse curviligne s et la coordonnée x. Nous pouvons alors décrire le rayon simplement par son altitude z en fonction de x. La seule composante non nulle de l'équation (4.34) est selon z et se met sous la forme:

$$\frac{d}{dx} n \frac{dz}{dx} = a n_0 \ . \tag{4.36}$$

En remarquant que n sort de la dérivation par rapport à x et en négligeant la variation de n par rapport à n_0 sur l'étendue verticale de la trajectoire (il serait à peine plus difficile d'en tenir compte), nous mettons l'équation du rayon sous la forme:

$$\frac{d^2 z}{dx^2} = a \ , \tag{4.37}$$

dont la solution est évidemment une parabole dont la concavité est dirigée vers le haut. Un rayon incident vers le sol en incidence rasante est réfléchi, ce qui donne au sol l'aspect d'un miroir à grande distance, d'où l'apparition de lacs au coeur des déserts....

Considérons enfin le problème de la fibre optique à gradient d'indice. Il existe des fibres à saut d'indice, dont le cœur, d'un diamètre de quelques microns, a un indice plus élevé que le reste de la fibre. La lumière est alors guidée par réflexion interne totale sur le dioptre limitant le cœur. Il existe aussi des fibres à gradient d'indice, où la variation d'indice est continue et en général parabolique en fonction de la distance à l'axe. Il faudrait évidemment, comme les dimensions transverses sont de l'ordre de la longueur d'onde, au moins pour les fibres monomodes, recourir à une approche ondulatoire. Nous allons nous contenter ici de comprendre qualitativement le confinement de la lumière par une image géométrique. Nous postulerons donc que l'indice est à une distance r de l'axe de la forme $n = n_0(1 - ar^2)$ $(a > 0)$. Nous emploierons une géométrie très voisine de celle du mirage. Dans un plan xOz (x est l'axe de la fibre), nous écrivons la trajectoire d'un rayon paraxial. Nous pouvons encore confondre s et x, remarquer que n est indépendant de x et varie très peu par rapport à sa valeur n_0 sur l'axe. Le rayon est alors entièrement décrit par $z(x)$ et son équation de propagation s'écrit:

$$\frac{d^2 z}{dx^2} = -2az \ , \tag{4.38}$$

dont la solution est évidemment une sinusoïde. Le rayon oscille donc autour de l'axe dont il reste à distance finie. Les rayons sont effectivement guidés dans cette géométrie. On remarquera que, si a était négatif, on obtiendrait des solutions divergentes où le rayon s'éloigne indéfiniment de l'axe. Comme on pouvait s'y attendre, une fibre ne guide la lumière que si l'indice est maximal au centre.

4.3.3 Principe de Fermat

Nous avons ici dérivé l'équation de propagation des rayons lumineux de l'approche eikonale. Elle peut en fait être déduite aussi du principe de Fermat, principe variationnel qui stipule que le trajet effectivement suivi par la lumière réalise un extremum du temps de parcours (en général un minimum). Montrons que c'est effectivement le cas par un calcul élémentaire de variations (on pourra se reporter aux résultats de la première et de la deuxième partie pour se familiariser à nouveau avec ces calculs dont nous ne détaillerons pas ici le principe).

Nous considérons un rayon lumineux décrit par l'équation intrinsèque $\mathbf{r}(s)$ et un rayon infiniment proche, partant du même point et arrivant au même point. L'écart entre ces deux rayons peut s'écrire $\delta\mathbf{r}(s)$, paramétré par l'abscisse curviligne sur le rayon de référence. Le temps de parcours sur le rayon de référence, $\int nds$, est extrémal. Sa variation au premier ordre dans les $\delta\mathbf{r}$ quand on passe au rayon varié doit être nulle. Or,

$$\delta \int nds = \int (\delta n) ds + \int n\delta ds \ , \tag{4.39}$$

où le premier terme décrit la variation d'indice quand on passe d'une trajectoire à l'autre, le second la variation de longueur de l'élément de trajectoire. On a évidemment:

$$\delta n = \boldsymbol{\nabla} n \cdot \delta\mathbf{r} \ . \tag{4.40}$$

Pour estimer le second terme, nous considérons un élément différentiel ds de la trajectoire de référence. L'élément correspondant de la trajectoire variée a pour longueur:

$$ds + \delta ds = \sqrt{ds^2 + 2d\mathbf{r} \cdot d\delta\mathbf{r}} \ , \tag{4.41}$$

d'où on tire sans peine:

$$\delta ds = \frac{d\mathbf{r}}{ds} d\delta\mathbf{r} \ . \tag{4.42}$$

Dans tout ce calcul, nous avons considéré les accroissements de trajectoire comme des infiniments petits par rapport aux éléments différentiels de ces mêmes trajectoires. En écrivant finalement l'intégration par parties traditionnelles:

$$\int n \frac{d\mathbf{r}}{ds} d\delta\mathbf{r} = -\int \frac{d}{ds} n \frac{d\mathbf{r}}{ds} \delta\mathbf{r} \ , \tag{4.43}$$

où le terme tout intégré est nul puisque trajectoire de référence et trajectoire variée coïncident aux extrémités, on met en facteur $\delta\mathbf{r}$. L'intégrale de variation n'est nulle pour tout accroissement que si le facteur de $\delta\mathbf{r}$ est nul. Il est évident que l'équation ainsi obtenue est précisément (4.34). L'approche eikonale est donc strictement équivalente au principe de Fermat.

Partie IV

Sources du champ électromagnétique

Introduction

Nous allons nous intéresser, dans cette partie, au champ électromagnétique produit par un certain nombre de sources. Si la solution des équations de Maxwell en termes de potentiels retardés et de fonctions de Green est parfaitement générale et explicite, elle ne fournit pas en général une solution physiquement transparente. Nous allons donc, dans cette partie, considérer des mouvements de charges ou des répartitions de courant bien particulières mais choisies pour leur grande importance physique, et écrire explicitement le champ produit. Très souvent, nous nous contenterons d'examiner les champs rayonnés à grande distance. Nous verrons en effet que le champ peut se mettre sous la forme d'un développement en puissances de l'inverse de la distance à la source r. A très grande distance, ce développement est dominé par le premier terme, en $1/r$. Des champs en $1/r$ conduisent à un flux d'énergie électromagnétique en $1/r^2$ et donc à un flux d'énergie fini à travers une sphère de rayon r. Seule cette partie du développement du champ conduit à une propagation à grande distance de l'énergie électromagnétique.

Nous examinerons essentiellement trois types de sources dans les trois chapitres de cette partie. Nous commencerons par étudier le champ rayonné par une particule unique, éventuellement relativiste, dont le mouvement est imposé mais arbitraire. Nous montrerons ainsi qu'une particule relativiste rayonne de façon particulièrement intense quand elle est accélérée et nous décrirons les effets et les applications de ce rayonnement qui joue un rôle de premier plan dans les accélérateurs de particules. Nous examinerons ensuite le cas beaucoup plus simple d'une particule ayant un mouvement sinusoïdal de petite amplitude autour de l'origine. Nous retrouverons ainsi le champ d'un dipôle électromagnétique oscillant. En raison de la très grande importance de ce problème dans des domaines variés, nous examinerons en détail le champ émis sans nous cantonner au champ rayonné.

Dans le deuxième chapitre, nous nous pencherons sur le champ rayonné par une répartition localisée de courants oscillants. Nous montrerons que le champ rayonné à une distance grande devant la taille caractéristique de la source peut s'exprimer comme un développement en puissances de l'extension caractéristique de la source divisée par la longueur d'onde émise. Le premier terme de ce "développement multipolaire" nous redonnera le rayonnement du dipôle que nous placerons ainsi dans un cadre plus général. Notons que les résultats obtenus ici seront au centre de notre approche de l'électromagnétisme dans la matière dans la prochaine partie. Nous terminerons ce chapitre en discutant de plusieurs répartitions de courants exemplaires. Nous aborderons en particulier très brièvement le problème des antennes, d'une importance cruciale pour les télécommunications.

Dans le troisième chapitre, nous appliquerons les résultats sur les rayonnements multipolaires au cas des sources microscopiques. Nous nous poserons en fait essentiellement le problème de la diffusion du rayonnement par un atome unique. Nous commencerons par donner un modèle classique, extrêmement naïf de prime abord, de la dynamique de l'atome. Nous pourrons alors explicitement calculer la réponse atomique à un champ incident. Nous essaierons ensuite de donner les éléments essentiels d'un modèle quantique de la réponse atomique à un champ classique (nous ne pourrons bien sûr donner le seul modèle rigoureux qui impose de quantifier le champ). Nous montrerons que les résultats du modèle classique sont étonnamment exacts pour la diffusion d'un champ faible non résonnant avec une fréquence de Bohr de l'atome. Nous conclurons ce chapitre en examinant très brièvement, dans le modèle quantique, la réponse d'un atome à un champ résonnant.

Chapitre 1

Rayonnement d'une particule en mouvement

Nous nous pencherons dans ce chapitre sur le problème du rayonnement d'un charge ponctuelle unique en mouvement arbitraire, éventuellement relativiste. Le problème général du mouvement d'une charge est extrêmement complexe. Il résulte en effet à la fois des forces imposées à la charge et du champ produit par la charge elle–même. Comme nous le verrons, une charge en mouvement rayonne. Ce rayonnement conduit à une perte d'énergie mécanique que l'on peut interpréter comme résultant d'une force produite sur la charge par son propre champ, la force de réaction de rayonnement. D'une part, la modélisation de cette force n'est pas sans poser des difficultés mathématiques (dans le cas d'une charge ponctuelle, on peut s'attendre à des problèmes de divergence — nous aborderons ce problème à la fin de ce chapitre). D'autre part, la résolution des équations est très difficile dans le cas général.

Nous nous cantonnerons donc à étudier le rayonnement d'une particule dont le mouvement est imposé. Nous négligerons l'action des forces de réaction de rayonnement par rapport aux forces qui fixent la trajectoire de la particule. Nous verrons, en discutant de situations physiques concrètes, que cette approximation est souvent excellente. Comme nous envisagerons le cas de particules relativistes, il serait tout naturel de nous placer dans le formalisme de l'électrodynamique relativiste et de n'écrire que des quantités explicitement covariantes, indépendantes du référentiel. Si cette approche est la plus élégante et, du point de vue des calculs, la plus compacte (on consultera à ce sujet le Jackson), elle ne révèle pas les interprétations physiques les plus transparentes des résultats obtenus. Nous verrons en particulier que le renforcement considérable du rayonnement dans la direction de la vitesse a une interprétation cinématique très simple. Nous nous placerons donc dans un référentiel $\mathcal{R}$ particulier et manipulerons l'espace et le temps dans ce référentiel uniquement. Bien sûr, les résultats finaux seront rigoureux. On pourra trouver les champs dans un référentiel quelconque par une simple transformation de Lorentz. Nous donnerons également pour mémoire les solutions explicitement covariantes.

1.1 Potentiels de Liénard–Wiechert

Nous considérons donc une particule de charge q dont le mouvement, imposé, est décrit par la trajectoire $\mathbf{r}_0(t')$ (voir figure 1.1). Nous cherchons le champ observé au point $\mathbf{r}$ à l'instant t. Pour cela, nous déterminerons d'abord les potentiels scalaires et vecteurs que nous dériverons ensuite pour obtenir les champs. Avant cette dérivation, dans tout le calcul des potentiels, la position "d'observation" $\mathbf{r}$ et l'instant t sont donc fixés. Nous noterons $\mathbf{n}(t')$ le vecteur unitaire de la particule, prise à l'instant t', vers le point d'observation. Nous noterons enfin $R(t') = |\mathbf{r} - \mathbf{r}_0(t')|$ la distance entre la particule et le point d'observation.

La densité de charge et de courant représentant la charge ponctuelle peuvent s'écrire:

$$\rho(\mathbf{r}', t') \quad = \quad q\delta(\mathbf{r}' - \mathbf{r}_0(t')) \tag{1.1}$$

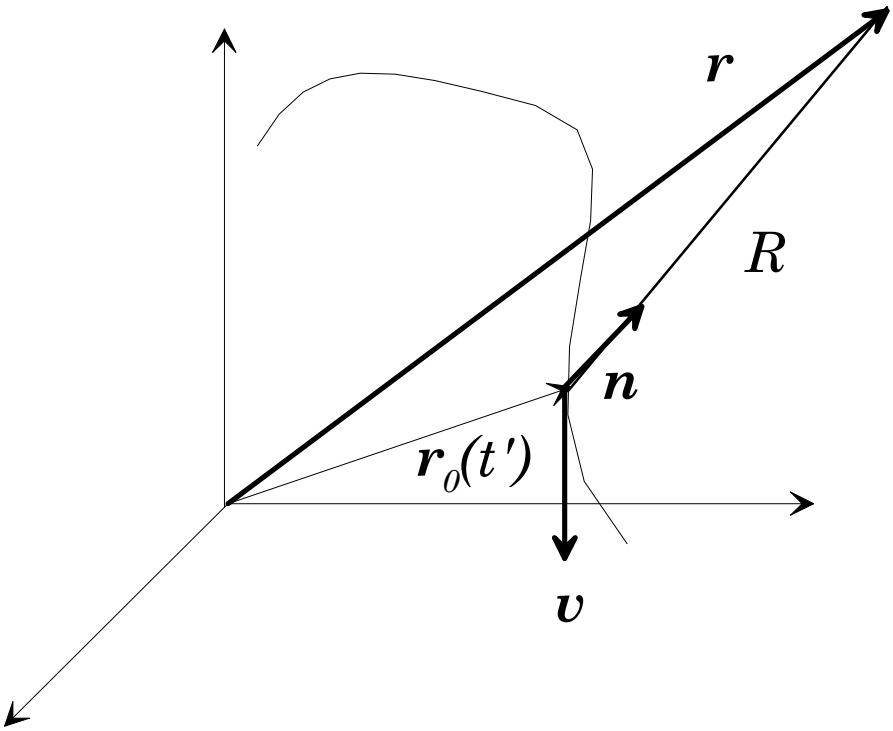

Figure 1.1: Rayonnement d'un charge en mouvement: notations.

$$\mathbf{j}(\mathbf{r}', t') \quad = \quad q\mathbf{v}(t')\delta(\mathbf{r}' - \mathbf{r}_0(t')) \ , \tag{1.2}$$

où $\mathbf{v} = d\mathbf{r}_0/dt'$ est la vitesse de la particule à l'instant t'. En nous plaçant en jauge de Lorentz, nous pouvons écrire les potentiels en fonction de ces densités de charge et de courant au moyen de la solution en potentiels retardés. Comme les calculs sont identiques pour le potentiel scalaire et chaque composante du potentiel vecteur, nous ne traiterons explicitement que le premier. Le potentiel peut donc s'écrire:

$$V(\mathbf{r}, t) = \frac{1}{4\pi\epsilon_0} \int d\mathbf{r}' dt' \frac{\rho(\mathbf{r}', t')\delta(t' - t + |\mathbf{r} - \mathbf{r}'(t')|/c)}{|\mathbf{r} - \mathbf{r}'(t')|} \ , \tag{1.3}$$

Où l'on notera que nous avons changé le signe de l'argument de la fonction δ spatio–temporelle par rapport aux conventions utilisées dans la partie précédente. Ce changement de signe, effectué pour des raisons de commodité, n'a bien sûr aucune conséquence sur le résultat final. On a donc encore:

$$V(\mathbf{r}, t) = \frac{q}{4\pi\epsilon_0} \int d\mathbf{r}' dt' \frac{\delta(\mathbf{r}' - \mathbf{r}_0(t'))\delta(t' - t + |\mathbf{r} - \mathbf{r}'(t')|/c)}{|\mathbf{r} - \mathbf{r}'(t')|} \ . \tag{1.4}$$

Sous cette forme, l'intégrale sur $\mathbf{r}'$ est triviale: seul contribue le point $\mathbf{r}' = \mathbf{r}_0(t')$. En posant alors:

$$g(t') = t' - t + \frac{R(t')}{c} \ , \tag{1.5}$$

on met V sous la forme:

$$V(\mathbf{r}, t) = \frac{q}{4\pi\epsilon_0} \int \frac{\delta(g(t'))}{R(t')} dt' \ . \tag{1.6}$$

Seul contribuera à l'intégrale sur le temps l'instant t_0 tel que $g(t_0)$ s'annule (en supposant, ce que nous allons montrer, que cet instant existe et est unique).

On peut facilement comprendre la signification physique de l'instant t_0. Il est tel que:

$$t_0 = t - \frac{R(t_0)}{c} \tag{1.7}$$

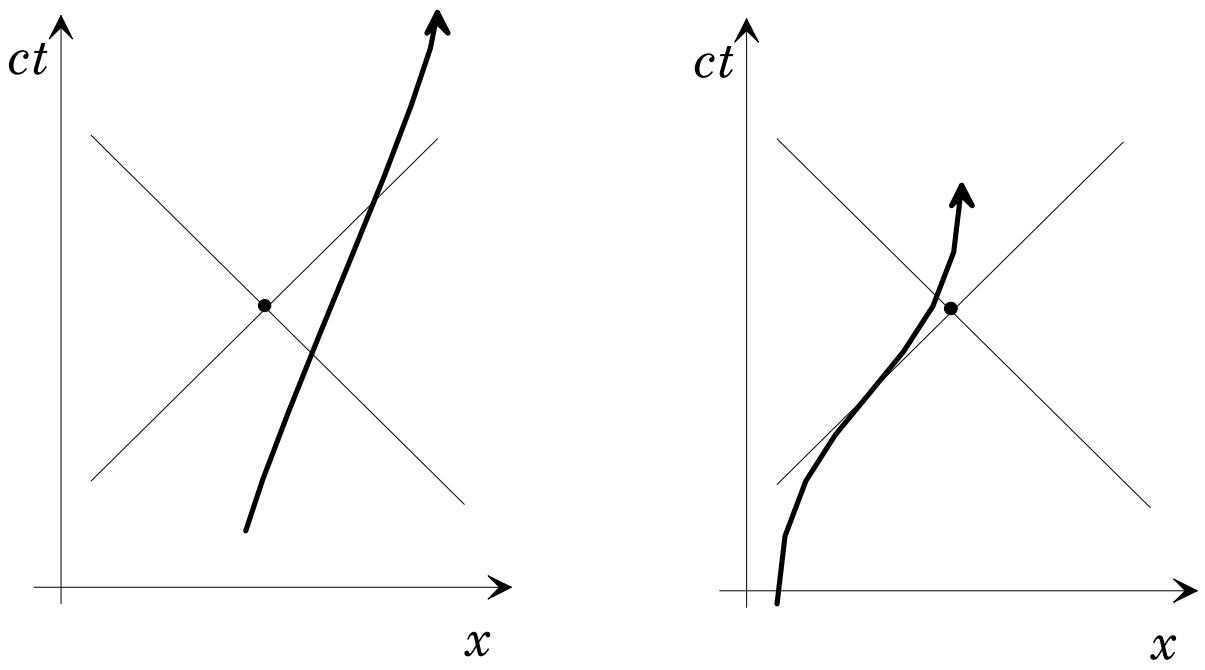

Figure 1.2: Intersection du cône de lumière de l'observateur et de la ligne d'univers de la particule. A gauche: particule s'éloignant de l'observateur. A droite: particule relativiste venant vers l'observateur. Seule compte pour le calcul des champs l'intersection située dans le passé. Le champ reçu est d'autant plus grand que la particule reste plus longtemps au voisinage du cône de lumière.

c'est à dire tel qu'un signal lumineux émis par la particule à l'instant t_0 arrive, après s'être propagé à la vitesse c, au point d'observation précisément à l'instant t. Bien sûr, la forme même de l'équation impose à t_0 d'être plus petit que t. En fait, t_0 est la coordonnée temporelle de l'intersection de la ligne d'univers de la particule avec le cône de lumière de l'événement d'observation (voir figure 1.2). Comme la pente de la ligne d'univers dans un diagramme d'espace–temps est toujours plus petite que celle du cône (la vitesse de toute particule chargée, donc massive, est toujours plus petite que c), il y a deux intersections et deux seulement entre la ligne d'univers et le cône de lumière. L'une est située dans le passé et sa coordonnée temporelle correspond à l'instant t_0 que nous cherchons. L'autre est située dans le futur de l'observateur et ne peut correspondre à une solution de l'équation donnant t_0. En fait, ce point dans le futur serait le seul à contribuer au potentiel si nous avions, en dépit de la causalité, choisi la solution en termes de potentiels avancés. Nous avons donc bien montré que l'instant t_0, qui va jouer un rôle central dans toute la suite de la discussion, existe et est unique.

L'équation (1.6) fait intervenir la distribution $\delta(g(t'))$. Pour l'identifier, faisons–la agir sur une fonction régulière arbitraire $f(t')$. En posant $u = g(t')$ (ce changement de variable est régulier puisque seuls les points voisins de t_0 contribuent à l'intégrale), nous pouvons écrire:

$$\int \delta(g(t'))f(t')dt' = \int \frac{\delta(u)f(g^{-1}(u))}{|g'(g^{-1}(u))|}du = \frac{f(t_0)}{|g'(t_0)|} \ , \tag{1.8}$$

en désignant par g' la dérivée de g par rapport à son argument. La valeur absolue de la dérivée au dénominateur est obtenue en ordonnant les bornes de l'intégration quand la fonction g est décroissante au voisinage de t_0. En termes de distributions, on a donc:

$$\delta(g(t')) = \frac{\delta(t' - t_0)}{|g'(t_0)|} \ . \tag{1.9}$$

Dans toute la suite, nous ne considérerons que des fonctions g qui sont croissantes au voisinage de leur zéro. Nous pourrons donc omettre à partir de maintenant les valeurs absolues au dénominateur. L'intégrale sur t' dans l'expression du potentiel V est donc triviale et on trouve:

$$V = \frac{q}{4\pi\epsilon_0}\frac{1}{g'(t_0)}\frac{1}{R(t_0)} \ . \tag{1.10}$$

Pour obtenir une forme plus explicite du potentiel, il nous reste à exprimer $g'(t_0)$ en fonction des paramètres cinématiques de la particule. A partir de (1.5), nous trouvons par simple dérivation:

$$g'(t') = 1 + \frac{\partial R(t')}{\partial t'} = 1 + \frac{\mathbf{R}}{R}\cdot\frac{d\mathbf{R}}{dt'} = 1 - \frac{\mathbf{n}(t')\cdot\mathbf{v}(t')}{c} \ , \tag{1.11}$$

où nous avons posé $\mathbf{R} = \mathbf{r} - \mathbf{r}_0$, dont la dérivée par rapport au temps t' est l'opposé de la vitesse de la particule en t'. On a donc $g'(t_0) = 1 - \mathbf{n}\cdot\mathbf{v}/c$. Nous conviendrons dans toute la suite de prendre les fonctions de t' à l'instant t_0 quand leur argument est omis. On a donc finalement:

$$V = \frac{q}{4\pi\epsilon_0}\frac{1}{1-\mathbf{n}\cdot\mathbf{v}/c}\frac{1}{R(t_0)} \tag{1.12}$$

et, par un raisonnement analogue,

$$\mathbf{A} = \frac{q}{4\pi\epsilon_0 c^2}\frac{1}{1-\mathbf{n}\cdot\mathbf{v}/c}\frac{\mathbf{v}}{R(t_0)} \ . \tag{1.13}$$

Ces expressions constituent les potentiels de Liénard–Wiechert. Ils ne diffèrent des potentiels de l'électrostatique ou de la magnétostatique que par le terme cinématique $1/g'$. Pour une particule immobile, on retrouve bien les expressions statiques. Pour une particule en mouvement, en revanche, les potentiels de Liénard peuvent être très différents des potentiels statiques. Pour une particule relativiste en particulier ($v \simeq c$), le facteur g' s'annule pratiquement quand $\mathbf{n}$ et $\mathbf{v}$ sont de même direction. Le potentiel est donc beaucoup plus grand dans la direction du mouvement de la particule que dans la direction opposée (le facteur g' étant alors voisin de 2). Le champ d'une particule relativiste est donc concentré vers l'avant. Nous verrons dans le prochain paragraphe les importantes conséquences de cette concentration.

On peut comprendre le renforcement des potentiels vers l'avant en termes purement cinématiques. Considérons les deux lignes d'univers illustrées sur la figure 1.2. A gauche, la particule relativiste traverse le cône de lumière de l'événement d'observation pratiquement perpendiculairement. A droite, la particule se dirige vers l'observateur et sa ligne d'univers est pratiquement tangente au cône de lumière. Dans le second cas, la particule reste beaucoup plus longtemps "en vue" de l'observateur que dans le premier. En d'autres termes, un intervalle de temps dt pour l'observateur correspond à un intervalle de temps retardé dt' pour la particule d'autant plus long qu'elle se dirige vers lui avec une plus grande vitesse. En quelque sorte, c'est un effet de sillage. Quand un bateau se déplace à une vitesse proche de celle des ondes capillaires à la surface de l'eau, les ondes émises à des instants différents s'accumulent et se renforcent à l'avant. En revanche, les ondes s'espacent vers l'arrière d'avantage que si le bateau restait immobile. Le même genre de phénomène, purement cinématique, est à l'origine du phénomène qui nous intéresse ici.

L'analogie en termes de sillage nous permet d'ailleurs de discuter qualitativement d'un autre phénomène intéressant. Si le bateau se déplace à une vitesse inférieure à celle des ondes capillaires, il n'y a pas de sillage à proprement parler. Le sillage "en V" correspond au cas où la vitesse du bateau est supérieure à celle de ces ondes. Les ondes rayonnées par le bateau à des instants différents s'additionnent alors le long d'un cône (deux lignes dans cet espace à deux dimensions) dont l'ouverture dépend du rapport de la vitesse des ondes et du bateau. Ces lignes constituent en fait une onde de choc et sont très analogues aux caustiques de l'optique. A trois dimensions, le même phénomène

s'observe pour un avion en vol supersonique. Le passage de l'onde de choc sur un observateur est responsable du célèbre "bang" (double en raison de la structure complexe de l'avion dont différentes parties génèrent des ondes de choc différentes). En l'absence de toute non linéarité, pour un "bateau" ponctuel, l'amplitude des ondes capillaires devrait diverger sur cette ligne. Nous ne pouvons bien sûr obtenir le même phénomène ici puisque la particule se déplace nécessairement moins vite que la lumière. Ce ne serait pas le cas si nous considérions une particule relativiste en mouvement dans un milieu matériel présentant un indice de réfraction élevé. Rien n'empêche alors la particule de se déplacer plus vite que la lumière dans le milieu. Il se forme alors un véritable "sillage", un cône le long duquel le champ électromagnétique est considérablement renforcé. Ce phénomène est responsable de la célèbre "radiation Cherenkov". Une lumière bleue relativement intense est émise par des particules en mouvement relativiste dans un milieu matériel. Ce rayonnement se manifeste par exemple autour des réacteurs immergés dans des piscines d'eau lourde. Il est également utilisé pour réaliser des détecteurs de particules, l'énergie de ce rayonnement donnant une indication de l'énergie de la particule. Même si ce phénomène n'existe pas dans le vide, nous allons montrer que le renforcement du champ "vers l'avant" a des conséquences importantes.

Notons pour finir que nous avons établi ici les potentiels de Liénard par un raisonnement non explicitement covariant, dans un référentiel donné. On peut établir une version purement relativiste de ces potentiels en écrivant une version complètement covariante des potentiels retardés. Le lecteur intéressé trouvera la dérivation explicite dans le Jackson. Nous ne donnerons ici que la forme finale du 4–potentiel:

$$\mathcal{A}^\mu = \frac{q}{4\pi\epsilon_0 c^2} \frac{U^\mu(\tau_0)}{U^\nu(R_\nu - R_{0\nu}(\tau_0))} \ , \tag{1.14}$$

où les R_ν sont les coordonnées covariantes de l'événement d'observation, les $R_{0\nu}$ celles de la particule et τ_0 le temps propre de la particule à l'événement où elle croise le cône de lumière de l'événement d'observation. Notons que ce que nous gagnons en généralité et en covariance est largement perdu en clarté de l'interprétation physique. L'aspect purement cinématique du renforcement vers l'avant est loin d'être apparent sur ce genre d'expression.

1.2 Champs rayonnés

Nous calculerons dans ce paragraphe les champ rayonnés par la particule. Il nous suffit pour cela, au moins formellement, de dériver les expressions des potentiels de Liénard–Wiechert. En fait, nous aurons à faire face à deux difficultés. La première est que, sans préciser d'avantage le mouvement de la particule, nous ne connaissons le temps "retardé" t_0 que comme solution d'une équation implicite. Il nous faudra donc procéder avec soin pour obtenir quand même des expressions explicites pour les champs. L'autre point est qu'en procédant aux dérivations des potentiels en $1/R$, nous allons obtenir les champs comme une somme de termes faisant intervenir $1/R$ et ses dérivées successives $1/R^2$ et $1/R^3$. Si les calculs ne présentent aucune difficulté de principe, ils sont assez lourds. De plus il n'est généralement pas indispensable de garder ces termes.

On s'intéresse en effet surtout en pratique au champ produit par la particule à une grande distance. On peut donc ne garder, dans ce développement, que le terme en $1/R$, qui domine tous les autres. Une autre manière de le voir est de considérer le vecteur de Poynting. Les termes en $1/R$ dans $\mathbf{E}$ et $\mathbf{B}$ donnent un flux d'énergie en $1/R^2$. La quantité d'énergie traversant une sphère de rayon R est donc finie. Ces termes correspondent bien à un transport d'énergie vers l'infini, à un rayonnement. En revanche, la combinaison de termes d'ordre supérieur dans $\mathbf{E}$ et $\mathbf{B}$ donne des contributions au vecteur de Poynting décroissant plus vite que $1/R^2$. La quantité d'énergie traversant la sphère de rayon R décroît donc avec R, au moins comme $1/R$. Ces termes ne correspondent qu'à des "champs proches" (comprenant, au moins, les champs statiques en $1/R^2$), en général sans importance. Nous ne garderons donc, au cours de nos dérivations, que les termes dominants dans les développements en

puissances de $1/R$. Nous verrons, dans le dernier paragraphe de ce chapitre, ce que sont les résultats quand on conserve tous les termes dans le cas particulièrement simple du dipôle oscillant.

Notons également qu'on peut s'attendre à ce que le champ rayonné présente une structure d'onde plane. Si la source est harmonique, les potentiels doivent avoir une structure d'onde sphérique. Suffisamment loin de l'origine et dans un volume suffisamment local, le vecteur d'onde (radial) est pratiquement uniforme et l'onde sphérique a pratiquement une structure d'onde plane. $\mathbf{E}$ et $\mathbf{B}$ doivent donc être orthogonaux entre eux et à $\mathbf{n}$ et leurs modules doivent être dans un rapport c. Pour une source quelconque, la linéarité de l'électromagnétisme et les propriétés de la transformée de Fourier assurent que cette géométrie se conserve. Nous pourrions nous contenter de ne calculer que le champ électrique et en déduire le champ magnétique par ce genre d'arguments. Nous ferons cependant les deux calculs pour conforter notre confiance dans ces arguments qualitatifs.

Avant d'entreprendre le calcul des champs rayonnés, nous allons établir un certain nombre de formules utiles.

1.2.1 Dérivées utiles

Nous aurons besoin de dériver la fonction implicite t_0 par rapport aux coordonnées et à l'instant d'observation. En effet t_0 dépend de t mais aussi de $\mathbf{r}$ par l'intermédiaire de (1.7). A partir de cette équation, on trouve en effet:

$$\begin{aligned} \frac{\partial t_0}{\partial t} &= 1 - \frac{1}{c}\frac{\partial R(t_0)}{\partial t} \\ &= 1 - \frac{1}{c}\frac{\partial R}{\partial t_0}\frac{\partial t_0}{\partial t} \ . \end{aligned} \tag{1.15}$$

Nous avons déjà calculé la dérivée de R par rapport à t_0 pour trouver l'expression de g'. Elle vaut $-\mathbf{v}\cdot\mathbf{n}$. On a donc:

$$\frac{\partial t_0}{\partial t} = 1 + \frac{\mathbf{n}\cdot\mathbf{v}}{c}\frac{\partial t_0}{\partial t} \ , \tag{1.16}$$

ce qui nous donne immédiatement:

$$\frac{\partial t_0}{\partial t} = \frac{1}{g'(t_0)} \ . \tag{1.17}$$

Nous trouvons ainsi une justification à nos raisonnement qualitatifs sur le renforcement cinématique des potentiels vers l'avant. Plus g' est petit, plus à un intervalle de temps donné pour l'observateur correspond un grand intervalle de temps retardé pour la particule. A la limite d'un mouvement ultra–relativiste, l'observateur "accumule" dans un intervalle de temps extrêmement court tous les champs produits par la particule.

On peut utiliser la même démarche pour exprimer $\boldsymbol{\nabla} t_0$ (la dérivée s'entendant par rapport à $\mathbf{r}$). On a

$$\boldsymbol{\nabla} t_0 = -\frac{1}{c}\boldsymbol{\nabla} R \ . \tag{1.18}$$

Dans le gradient de R, nous devons distinguer deux termes. Le premier provient de la dépendance explicite de $R = |\mathbf{r} - \mathbf{r}_0|$ en $\mathbf{r}$. Le gradient associé est évidemment radial. Le deuxième provient de la dépendance implicite de R par l'intermédiaire de $\mathbf{r}_0$, qui est lui même une fonction de t_0 et donc de $\mathbf{r}$. De manière évidente, ces deux termes s'ajoutent et on a:

$$\boldsymbol{\nabla} t_0 = -\mathbf{n}\frac{\partial|\mathbf{r}-\mathbf{r}_0|}{c\partial r} - \frac{1}{c}\frac{\partial R}{\partial t_0}\boldsymbol{\nabla} t_0 \ . \tag{1.19}$$

La première dérivée partielle s'entend à $\mathbf{r}_0$ constant. Elle est évidemment égale à 1. La deuxième vient d'être rappelée. On trouve donc:

$$\boldsymbol{\nabla} t_0 = -\frac{\mathbf{n}}{c} + \frac{\mathbf{n}\cdot\mathbf{v}}{c}\boldsymbol{\nabla} t_0 \tag{1.20}$$

et finalement:

$$\boldsymbol{\nabla} t_0 = -\frac{1}{g'(t_0)}\frac{\mathbf{n}}{c} \ . \tag{1.21}$$

Nous avons pu mettre les dérivées utiles sous une forme particulièrement simple, toujours en termes du facteur cinématique $1/g'$.

1.2.2 Champ électrique

Le champ électrique s'écrit bien sûr

$$\mathbf{E} = -\frac{\partial \mathbf{A}}{\partial t} - \boldsymbol{\nabla} V \ . \tag{1.22}$$

La dérivée partielle de $\mathbf{A}$ par rapport au temps s'écrit:

$$-\frac{\partial \mathbf{A}}{\partial t} = -\frac{\partial \mathbf{A}}{\partial t_0}\frac{\partial t_0}{\partial t} = -\frac{1}{g'(t_0)}\frac{q}{4\pi\epsilon_0 c^2}\frac{\partial}{\partial t_0}\left[\frac{\mathbf{v}}{g'(t_0)R(t_0)}\right] \ . \tag{1.23}$$

Dans la dérivée du crochet par rapport au temps retardé t_0, il apparaît des termes en $1/R$ correspondant à la dérivée de $\mathbf{v}/g'$. La dérivée de $1/R$, pour sa part, donne des termes en $1/R^2$ que nous devons négliger. A cet ordre d'approximation, nous avons donc:

$$\begin{aligned} -\frac{\partial \mathbf{A}}{\partial t} &= -\frac{q}{4\pi\epsilon_0 c^2}\frac{1}{g'(t_0)}\frac{1}{R}\frac{\partial}{\partial t_0}\frac{\mathbf{v}}{g'(t_0)} \\ &= -\frac{q}{4\pi\epsilon_0 c^2}\frac{1}{g'^3(t_0)}\frac{1}{R}\left[\mathbf{a}g' - \mathbf{v}g''\right] \ , \end{aligned} \tag{1.24}$$

l'accélération $\mathbf{a}$ étant elle aussi évaluée au temps retardé t_0.

Le gradient du potentiel scalaire fait intervenir la dérivée partielle par rapport au temps retardé multipliée par le gradient du temps retardé par rapport au point d'observation. En toute rigueur, il fait aussi intervenir le gradient de la dépendance explicite du potentiel retardé par rapport au point d'observation. Ce gradient est toutefois en $1/R^2$, tant pour la dérivée de la partie électrostatique en $1/R$ que pour la dérivée du facteur cinématique. Celle ci fait en effet intervenir le gradient de $\mathbf{n}$, qui est en $1/R$, facteur qui est à multiplier par la dépendance électrostatique en $1/R$. A l'ordre d'approximation où nous nous sommes placés, ces termes sont négligeables et nous ne conserverons que le premier terme. On a donc:

$$-\boldsymbol{\nabla} V = -\frac{\partial V}{\partial t_0}\boldsymbol{\nabla} t_0 = \frac{q}{4\pi\epsilon_0}\frac{\mathbf{n}}{cg'}\frac{\partial}{\partial t_0}\left[\frac{1}{g'}\frac{1}{R}\right] \ . \tag{1.25}$$

Comme précédemment, nous ne garderons pas les termes faisant intervenir la dérivée de $1/R$ et nous avons:

$$-\boldsymbol{\nabla} V = -\frac{q}{4\pi\epsilon_0}\frac{\mathbf{n}g''}{cg'^3}\frac{1}{R} \ . \tag{1.26}$$

En regroupant avec les termes provenant du potentiel vecteur, nous trouvons finalement l'expression complète du champ électrique:

$$\mathbf{E} = -\frac{q}{4\pi\epsilon_0 c^2}\frac{1}{R}\frac{1}{g'^3}\left[(c\mathbf{n} - \mathbf{v})g'' + \mathbf{a}g'\right] \ . \tag{1.27}$$

Il ne nous reste plus à exprimer que g''. Partant de $g' = 1 - \mathbf{n}\cdot\mathbf{v}/c$, nous obtenons

$$g'' = -\frac{1}{c}\frac{\partial \mathbf{n}}{\partial t_0}\cdot\mathbf{v} - \frac{1}{c}\mathbf{n}\cdot\mathbf{a} \ . \tag{1.28}$$

La dérivée de $\mathbf{n}$ par rapport à t_0 est évidemment d'ordre $1/R$. Ce terme donnerait une contribution au champ en $1/R^2$ et doit être écarté. A l'ordre où nous nous plaçons, $g'' = -\mathbf{n}\cdot\mathbf{a}/c$. Notons

finalement qu'on peut regrouper les termes entre crochets dans l'expression du champ en un double produit vectoriel et écrire enfin:

$$\mathbf{E} = \frac{q}{4\pi\epsilon_0 c^2}\frac{1}{R}\frac{1}{(1-\mathbf{n}\cdot\mathbf{v}/c)^3}\mathbf{n}\times[(\mathbf{n}-\mathbf{v}/c)\times\mathbf{a}] \ . \tag{1.29}$$

Nous discuterons du contenu physique de cette équation après avoir calculé le champ magnétique.

1.2.3 Champ magnétique

Nous calculerons bien sûr le champ magnétique par $\mathbf{B} = \boldsymbol{\nabla}\times\mathbf{A}$. On a donc:

$$\mathbf{B} = \frac{q}{4\pi\epsilon_0 c^2}\boldsymbol{\nabla}\times\frac{\mathbf{v}}{g'R} \ . \tag{1.30}$$

Dans cette expression, les trois termes dans le rotationnel dépendent de la position, soit directement, soit par l'intermédiaire du temps retardé. Notons d'abord que dériver le terme en $1/R$ ferait apparaître des termes en $1/R^2$ que nous devons écarter. Nous pouvons donc écrire:

$$\mathbf{B} = \frac{q}{4\pi\epsilon_0 c^2}\frac{1}{R}\boldsymbol{\nabla}\times\frac{\mathbf{v}}{g'} \ . \tag{1.31}$$

Dans cette expression, en posant $\mathbf{u} = \mathbf{v}/g'$,

$$\begin{aligned}\boldsymbol{\nabla}\times\mathbf{u} &= e_{ijk}\partial_j u_k = e_{ijk}\partial_j t_0\frac{\partial u_k}{\partial t_0}\\ &= \boldsymbol{\nabla}t_0\times\frac{\partial\mathbf{u}}{\partial t_0} = \boldsymbol{\nabla}t_0\times(\frac{\mathbf{a}}{g'}-\frac{\mathbf{v}g''}{g'^2})\\ &= \frac{1}{g'^2}\frac{\mathbf{n}}{c}\times\mathbf{a}+\frac{g''}{g'^3}\frac{\mathbf{n}}{c}\times\mathbf{v} \ .\end{aligned} \tag{1.32}$$

Nous avons utilisé les conventions d'Einstein[1] et le tenseur complètement antisymétrique de rang 3, e_{ijk}.

Le champ magnétique s'écrit donc finalement:

$$\mathbf{B} = \frac{q}{4\pi\epsilon_0 c^2}\frac{1}{R}\frac{1}{g'^3}\frac{1}{c}\mathbf{n}\times\left[\mathbf{v}g''-\mathbf{a}g'\right] \ . \tag{1.33}$$

En comparant cette expression à (1.27), nous remarquons immédiatement que

$$\mathbf{B} = \frac{1}{c}\mathbf{n}\times\mathbf{E} \ . \tag{1.34}$$

1.2.4 Discussion physique

L'équation précédente indique, comme nous y attendions, que l'onde émise présente localement la structure d'une onde plane. Le champ se propage selon la direction radiale $\mathbf{n}$, les champs électrique et magnétique étant perpendiculaires et dans un rapport c. Ces champs ne sont non nuls que si la particule est accélérée. Une particule animée d'un mouvement de translation uniforme ne rayonne pas. C'est une simple conséquence de l'invariance relativiste. Si la particule rayonne, elle doit emprunter l'énergie émise à son énergie cinétique. Une particule qui rayonne doit voir sa vitesse se modifier. En se plaçant dans le référentiel initial de la particule, cela indiquerait qu'une charge au repos pourrait acquérir

[1] Encore une fois, dans l'espace à trois dimensions, il n'y a pas lieu alors d'établir de différence entre les coordonnées covariantes et contravariantes. Nous écrirons en conséquence tous les indices en position basse. Nous adopterons systématiquement cette convention dans la suite de ce chapitre.

spontanément une vitesse, ce qui est absurde. Notons enfin que le renforcement cinématique dans la direction du mouvement est considérablement accru pour les champs. Le facteur $1/g'$ apparaissant dans les potentiels est en effet remplacé par un facteur $1/g'^3$. Ce renforcement est simplement du au fait que les potentiels, tout en étant plus grands dans la direction du mouvement, ont aussi une dépendance spatiale plus rapide. Notons toutefois que les champs sont considérablement réduits dans la direction même du mouvement pour une particule ultra relativiste (dépendance en $\mathbf{n} - \mathbf{v}/c$ de l'équation (1.29)). Pour une particule animée d'une vitesse faible devant c, le champ est réparti dans tout l'espace et peut même être maximum dans une direction orthogonale à la vitesse (on le verra dans le cas du dipôle).

Il est facile de calculer, à partir des champs, le vecteur de Poynting décrivant l'énergie rayonnée. Notons d'abord qu'il est colinéaire à $\mathbf{n}$, dirigé vers l'extérieur. Son module s'écrit:

$$\Pi = \frac{q^2}{16\pi^2\epsilon_0 c^3} \frac{|\mathbf{n} \times [(\mathbf{n} - \mathbf{v}/c) \times \mathbf{a}]|^2}{R^2(1 - \mathbf{n} \cdot \mathbf{v}/c)^6} \ . \tag{1.35}$$

Le facteur cinématique est ici élevé à la puissance 6. Le rayonnement d'une particule ultra–relativiste s'effectue essentiellement vers l'avant en termes énergétiques. Cette expression doit être utilisée avec prudence si on désire déterminer la perte d'énergie de la particule à partir de l'énergie du champ rayonné. Nous avons calculé ici le flux d'énergie par unité de temps au point d'observation. Par intégration spatiale sur une sphère de rayon R centrée sur la particule, on en déduira l'énergie rayonnée par unité de temps dans le référentiel du laboratoire. Pour estimer l'effet du rayonnement sur le mouvement de la particule, il nous faudra estimer la quantité d'énergie émise par unité du temps retardé t_0. On prendra garde au facteur cinématique $1/g'$ entre ces deux unités de temps.

Pour des particules ultra–relativistes, l'énergie rayonnée peut être très importante. Ce rayonnement peut avoir des conséquences utiles ou néfastes. Il est d'abord utile comme source intense de rayonnement électromagnétique de haute fréquence. Le cône de rayonnement est en effet si étroit qu'il ne passe sur l'observateur, au cours du mouvement de la particule, que pendant un très bref instant. Les fréquences caractéristiques associées à une impulsion brève étant élevées, le rayonnement d'une particule relativiste peut atteindre le domaine des rayons X ou γ. Les tubes standard à rayons X, utilisés en radiographie, utilisent une forme "désordonnée" de ce rayonnement. Des électrons d'énergie modérée (50 keV) frappent une anode (appelée "anticathode" dans ce cas) constituée d'un métal lourd. Passant au voisinage de noyaux fortement chargés, les électrons sont fortement accélérés. Ils rayonnent alors leur énergie sous forme de lumière. En fait, ils perdent une fraction notable de leur énergie à chaque déflexion, sous forme de photons X dont l'énergie est voisine de l'énergie initiale de l'électron (notons que nos approximations ne sont guère valables dans ce cas). C'est en raison de ce mécanisme de perte d'énergie dans la matière que le rayonnement des particules accélérées est souvent appelé Bremsstrahlung, ou rayonnement de freinage. On comprend ainsi aussi pourquoi les électrons pénètrent beaucoup moins dans la matière que des particules plus lourdes (protons, par exemple). Etant beaucoup plus légers, ils sont plus fortement accélérés par les champs nucléaires et perdent leur énergie plus rapidement. En pratique, des électrons d'énergie modérée ne pénètrent que de quelques microns au plus dans de la matière dense, alors que des protons d'énergie comparable sont capables de l'irradier beaucoup plus profondément.

Le rayonnement de freinage dans les anticathodes des tubes à rayons X n'est pas du tout contrôlé. Le rayonnement émis est essentiellement isotrope et présente un très large spectre. Comme il est très difficile de réaliser des composants optiques dans le domaine des rayons X, il n'est guère possible de récupérer et de refocaliser ce rayonnement pour des applications fines. On doit souvent se contenter des "ombres chinoises" de la radiographie classique. On peut contrôler beaucoup mieux le rayonnement en organisant l'accélération. C'est par exemple ce qu'on réalise dans les sources de "rayonnement synchrotron" (nous verrons plus loin l'origine de ce nom) à "onduleur". On fait passer un faisceau d'électrons énergétiques dans un champ magnétique intense spatialement modulé (produit, par exemple, par une chaîne d'aimants permanents montés en quinconce). La trajectoire de l'électron est alors

ondulée, avec une forte accélération au voisinage des extrema de l'élongation transversale. Un rayonnement intense est alors émis dans la direction de propagation, avec un spectre relativement étroit. On peut ainsi réaliser des source intenses et directives de rayonnement X. Leur seul inconvénient est de nécessiter un anneau de stockage pour les électrons! Pour des intensités de faisceau suffisantes et dans le domaine micro-onde ou optique, le milieu électronique peut présenter un gain suffisant par émission stimulée pour le fonctionnement d'un laser. La fréquence peut en principe être accordée dans une très large gamme en modifiant l'énergie des électrons. Très utilisés dans le domaine infrarouge proche ou lointain, ces lasers n'ont été utilisés dans le domaine visible qu'à l'état de démonstrations de principe. Les lasers à colorant ou à sources solides, beaucoup plus simples et moins coûteux, ont en effet des performances supérieures.

Le rayonnement de freinage peut avoir des conséquences aussi pour des énergies beaucoup plus faibles. Dans le premier modèle quantique de Bohr Sommerfeld, les électrons orbitent autour du noyau sur des orbites elliptiques classiques. Certains paramètres de l'orbite sont quantifiés, c'est à dire déterminés par un nombre entier: l'énergie par le nombre quantique principal, le moment cinétique et l'orientation de l'orbite par les nombres quantiques orbitaux et magnétiques. Pour les orbites les plus elliptiques, l'électron est très fortement accéléré quand il passe près du noyau. Il doit donc rayonner une quantité d'énergie importante et on peut s'attendre à ce que la durée de vie de ces niveaux soit assez brève. Les orbites de moment angulaire maximum, au contraire, correspondent à une trajectoire électronique circulaire. L'accélération est minimale et on peut s'attendre à ce que la durée de vie de ces niveaux soit beaucoup plus importante. Ce raisonnement semi-classique ne peut pas rendre correctement compte des propriétés des niveaux profonds. En revanche, en vertu du principe de correspondance, il doit décrire convenablement les niveaux de grands nombres quantiques principaux, les "états de Rydberg". On constate en effet que les états de Rydberg "circulaires", de moment orbital maximum, ont une durée de vie plus de 100 fois plus longue que des états de même énergie, mais de faible moment angulaire. Nous donnerons les bases du calcul de ces durées de vie dans les paragraphes suivants.

Si le rayonnement de freinage a des effets bénéfiques, il est aussi une limitation importante aux performances des accélérateurs de particules. Dans les accélérateurs les plus répandus, descendants du "synchrotron", on courbe avec un champ magnétique la trajectoire des particules en forme de cercle. On peut ainsi accélérer les particules à chaque tour en les faisant passer dans une cavité contenant un champ radiofréquence. La fréquence est ajustée de telle manière que les particules rencontrent toujours dans la cavité un champ de même direction. Entre les zones de champ électrique, l'accélération des particules sur leur trajectoire circulaire fait qu'elles perdent de l'énergie par rayonnement. C'est là une limite sérieuse à l'efficacité de ces accélérateurs. Là encore, l'effet est beaucoup plus marqué pour des électrons que pour des particules lourdes. On peut le réduire en augmentant le rayon du cercle. C'est ainsi qu'on ne peut atteindre, au CERN, des énergies électroniques de 50 ou 100 GeV qu'au prix d'un anneau de 27 km de périmètre. La perte d'énergie par rayonnement, qui doit être compensée par les cavités accélératrices, est de quelques dizaines de MW, correspondant à quelques kW par mètre linéaire de faisceau. Ce rayonnement étant essentiellement émis dans le domaine des X durs, on comprend qu'il soit nécessaire d'enterrer profondément l'anneau. Le champ magnétique nécessaire pour obtenir un tel rayon est relativement modéré, de l'ordre de 1000 gauss (0.1T). Ce n'est pas du tout un facteur limitatif. Pour des accélérateurs à protons, en revanche, la limitation à l'énergie vient plus du champ magnétique nécessaire à la courbure des trajectoires que du rayonnement de freinage. Notons enfin que le rayonnement de freinage est beaucoup moins important dans les accélérateurs linéaires (parce que vitesse et accélération sont colinéaires). L'énergie est alors limitée par le champ électrique maximum (10MV/m) et la longueur du dispositif (1 km environ).

1.3 Réaction de rayonnement

Nous avons fait l'approximation, dans tout ce chapitre, que le rayonnement ne modifiait pas le mouvement de la charge. Il est bien évident que la puissance rayonnée doit être prise sur l'énergie mécanique de la particule et qu'on ne peut négliger cette réaction du rayonnement sur la trajectoire qu'en première approximation. Nous allons tenter dans ce paragraphe de décrire la perturbation au mouvement due au rayonnement. Pour simplifier la discussion, nous choisirons de ne travailler qu'avec un électron. Nous verrons en effet que la réaction de rayonnement a des conséquences sur le mouvement d'autant plus importantes que la masse de la particule est petite. L'adaptation de ces discussions à d'autres particules ne présenterait aucune difficulté. Nous ne nous placerons que dans un cadre non relativiste. Nous allons commencer, dans une discussion qualitative, par donner quelques ordres de grandeur. Nous pourrons ainsi dégager une échelle naturelle de temps pour ces phénomènes. Nous montrerons que la perturbation est en effet petite tant que les constantes caractéristiques du mouvement sont longues à cette échelle. Nous montrerons ensuite que l'action du rayonnement sur l'énergie de la particule peut se décrire comme le travail d'une force, la "force de réaction de rayonnement". Le premier appendice de ce chapitre montrera comment on peut obtenir l'expression de cette force en se donnant un modèle naïf de particule chargée et en estimant l'action sur le mouvement de la particule du champ produit par la particule elle–même. Nous donnerons alors quelques applications de la force de réaction de rayonnement. Nous montrerons en particulier comment elle donne une durée de vie finie (et très brève) aux atomes dans un modèle planétaire classique.

1.3.1 Approche qualitative

Pour un mouvement non relativiste, on peut, dans l'expression (1.35) du vecteur de Poynting des champs de Liénart Wiechert, négliger les termes en $\mathbf{v}/c$. On a alors:

$$\mathbf{\Pi} = \frac{q^2}{16\pi^2\epsilon_0 c^3}\frac{|\mathbf{n}\times(\mathbf{n}\times\mathbf{a})|^2}{R^2} \ . \tag{1.36}$$

En prenant l'axe Oz dans la direction de $\mathbf{a}$, on voit que $\mathbf{n}\times(\mathbf{n}\times\mathbf{a}) = a\sin\theta\mathbf{u}_\theta$ (avec les notations standard des coordonnées sphériques). L'intégration du vecteur de Poynting sur une sphère de rayon R ne pose alors aucune difficulté (tous les points de cette sphère "voient" manifestement la particule au même instant retardé et donc avec la même accélération). On met alors la puissance totale rayonnée par la charge sous la forme (formule de Larmor):

$$\mathcal{P} = \frac{q^2a^2}{6\pi\epsilon_0 c^3} \ . \tag{1.37}$$

Estimons maintenant l'effet de cette puissance sur le mouvement. Considérons pour cela d'abord une particule initialement au repos accélérée uniformément pendant une période T. L'énergie totale rayonnée pendant l'accélération s'écrit alors $\mathcal{P}T$. L'énergie mécanique finale de la charge est, pour sa part, égale à $ma^2T^2/2$. Le rayonnement aura une influence faible sur le mouvement si $\mathcal{P}T \ll ma^2T^2/2$. On peut aussi écrire cette condition sous la forme:

$$T \gg \tau \ , \tag{1.38}$$

avec

$$\tau = \frac{1}{6\pi\epsilon_0}\frac{q^2}{mc^3} \ , \tag{1.39}$$

la puissance rayonnée étant alors simplement $\mathcal{P} = m\tau a^2$. Le temps caractéristique τ, qui va jouer un rôle essentiel dans la suite de cette discussion, est maximal pour l'électron, la plus légère des particules chargées. Sa valeur numérique est de $6.32\,10^{-24}$ s. Notons que $c\tau$ est de l'ordre du "rayon classique"

r_e de l'électron. Ce rayon est donné par un modèle électrostatique très naïf de l'électron, sous la forme d'une sphère uniformément chargée en surface, de rayon r_e et portant la charge élémentaire. En identifiant l'énergie électrique de cette distribution avec l'énergie de masse de l'électron, on fixe la valeur du rayon. La valeur obtenue est, à un coefficient près (2/3 en l'occurrence), la quantité $c\tau$. La très faible valeur de τ indique que l'effet du rayonnement est tout à fait négligeable, sauf si on s'intéresse à des périodes d'accélération infiniment courtes.

On peut préciser encore cette approche qualitative en considérant une particule chargée animée d'un mouvement sinusoïdal de fréquence ω_0. L'accélération est alors de l'ordre de $\omega_0^2 d$ où d est l'extension du mouvement. L'énergie mécanique est pour sa part de l'ordre de $m\omega_0^2 d^2$. La puissance rayonnée pendant une période, $\mathcal{P}/\omega_0$, sera petite par rapport à l'énergie mécanique et donc le rayonnement sera une petite perturbation au mouvement, si:

$$\frac{1}{6\pi\epsilon_0}\frac{q^2}{c^3\omega_0}(\omega_0^2 d)^2 \ll m\omega_0^2 d^2 \ , \tag{1.40}$$

c'est à dire si:

$$\eta = \omega_0\tau \ll 1 \ . \tag{1.41}$$

Là encore, cette condition est très bien vérifiée pour tout mouvement à une fréquence raisonnable. L'effet du rayonnement sur le mouvement est donc une faible perturbation et il est bien justifié de considérer, au moins sur un intervalle de temps assez bref, que le mouvement de la particule est imposé.

Si le rayonnement ne modifie pas dramatiquement le mouvement, il n'en reste pas moins que, sur des temps longs, l'effet cumulatif de la puissance rayonnée finira par amortir le mouvement dont l'énergie mécanique sera lentement décroissante. Pour tenir compte plus commodément de cet effet, nous allons monter que la puissance rayonnée peut être modélisée par le travail d'une force.

1.3.2 Force de réaction de rayonnement

Considérons donc une particule en mouvement imposé entre les instants t_1 et t_2. En exprimant la puissance rayonnée en fonction de la vitesse de la particule, on peut écrire la variation de l'énergie de la particule due au rayonnement pendant cette période, $\Delta\mathcal{E}$, sous la forme:

$$\Delta\mathcal{E} = -\frac{1}{6\pi\epsilon_0}\frac{q^2}{c^3}\int_{t_1}^{t_2} \dot{\mathbf{v}}^2\, dt \ . \tag{1.42}$$

L'intégrale peut aisément être transformée par une intégration par parties. On a:

$$\int_{t_1}^{t_2} \dot{\mathbf{v}}^2\, dt = -\int \ddot{\mathbf{v}}\cdot\mathbf{v}\, dt + [\dot{\mathbf{v}}\cdot\mathbf{v}]_{t_1}^{t_2} \ . \tag{1.43}$$

Si la période considérée est suffisamment longue et si les vitesses et accélérations sont bornées, l'intégrale dans le second membre domine le terme tout intégré (à condition bien sûr qu'elle soit non nulle - nous excluons donc de fait le cas du mouvement uniformément accéléré). On peut donc écrire l'énergie $\Delta\mathcal{E}$ sous la forme du travail d'une force:

$$\mathcal{E} = \int \mathbf{F}_r\cdot\mathbf{v}\, dt \ , \tag{1.44}$$

où la "force de réaction de rayonnement" $\mathbf{F}_r$ est définie par:

$$\mathbf{F}_r = \frac{q^2}{6\pi\epsilon_0 c^3}\dot{\mathbf{a}} = m\tau\dot{\mathbf{a}} \ . \tag{1.45}$$

Cette force, proportionnelle à la dérivée temporelle de l'accélération, est tout à fait particulière et doit être utilisée avec beaucoup de précautions. Elle conduit tout d'abord, évidemment, à des équations du

mouvement qui sont du troisième ordre par rapport au temps. Il faut donc, pour spécifier complètement le mouvement, donner non seulement la position et la vitesse, mais aussi l'accélération initiale de la particule. Un tel comportement est étranger à tous les principes de la mécanique Newtonienne.

Un autre problème grave relatif à cette force est la présence, parmi les solutions aux équations du mouvement, de solutions divergentes, même pour une particule libre. Considérons en effet une particule initialement au repos. L'équation du mouvement s'écrit:

$$m\dot{\mathbf{v}} = m\tau\ddot{\mathbf{v}} \ , \tag{1.46}$$

dont la solution générique s'écrit:

$$\dot{\mathbf{v}} = \mathbf{a}_0 e^{t/\tau} \ . \tag{1.47}$$

Si on ne spécifie pas que l'accélération de la particule est initialement nulle, l'accélération et la vitesse, croissent exponentiellement avec une constante de temps, τ, extrêmement brève. Ces solutions sont complètement non physiques.

Mentionnons enfin une dernière difficulté avec la force de réaction de rayonnement. Elle prédit une force nulle pour un mouvement uniformément accéléré, donc a priori une énergie rayonnée nulle. Mais, pour ce type de mouvement, la puissance prédite par la formule de Larmor, est non nulle. On peut, bien sûr, constater que nos hypothèses pour l'intégration par parties conduisant à l'expression de $\mathbf{F}_r$ ne sont pas valables dans ce cas. Il faut retenir de cette discussion que la force de réaction de rayonnement est un intermédiaire commode pour estimer l'influence du rayonnement sur le mouvement, mais qu'elle doit être utilisée avec précaution, le plus souvent dans une approche perturbative.

Dans l'appendice à ce chapitre, nous présenterons un calcul, dû à Abraham et Lorentz, qui justifie dans une certaine mesure la forme de la force de réaction de rayonnement. On y considère la charge comme une répartition étendue de densité de charge et on estime la force qu'elle exerce sur elle-même. Après un calcul fastidieux, on trouve une contribution proportionnelle à l'accélération, qu'on peut interpréter comme la masse de la particule, et la contribution de la force de réaction de rayonnement. Comme on le verra dans l'appendice, ce modèle présente des inconsistances graves et doit aussi être pris avec précautions.

1.3.3 Application

Nous allons utiliser, avec précautions, la force de réaction de rayonnement pour estimer la durée de vie des atomes dans le modèle de Rutherford (nous discuterons à nouveau ce problème en détail dans le chapitre sur le rayonnement des sources atomiques). La réaction de rayonnement, en réduisant graduellement l'énergie de l'électron en orbite autour du noyau, doit le conduire à tomber sur le noyau. Pour simplifier les calculs, nous considérerons un atome d'hydrogène et nous supposerons que les orbites sont circulaires, comparables à celles du modèle de Bohr. Nous nous servirons d'ailleurs des prédictions de ce modèle pour fixer les ordres de grandeur relatifs à nos orbites (en particulier celui de leur rayon).

Nous supposerons que l'énergie mécanique $\mathcal{E}$ évolue lentement par rapport au mouvement orbital (nous faisons ici une approximation séculaire comparable à celles qui sont utilisées en astronomie). Nous noterons $U(\mathbf{r}) = U(r) = \alpha/r$ le potentiel Coulombien à symétrie sphérique, avec

$$\alpha = \frac{q^2}{4\pi\epsilon_0} \ . \tag{1.48}$$

La dérivée temporelle de l'énergie mécanique est donnée par le travail de la force de réaction de rayonnement, ou, plus simplement encore, par la formule de Lorentz (on peut remarquer que, pour un mouvement périodique, le terme tout intégré dans le raisonnement conduisant à $\mathbf{F}_r$ est identiquement nul à condition d'intégrer sur une période; la formule de Larmor est donc strictement équivalente au travail de $\mathbf{F}_r$). On a alors:

$$\frac{d\mathcal{E}}{dt} = -m\tau\overline{\mathbf{a}^2} \ , \tag{1.49}$$

où la barre désigne une moyenne temporelle sur une période. En remarquant que l'accélération est largement dominée par la force coulombienne et peut donc s'écrire:

$$\mathbf{a} = -\frac{1}{m}\frac{dU}{dr}\mathbf{u}_r \ , \tag{1.50}$$

on obtient finalement:

$$\frac{d\mathcal{E}}{dt} = -\frac{\tau}{m}\overline{\left(\frac{dU}{dr}\right)^2} \ . \tag{1.51}$$

On notera qu'on pourrait écrire une équation similaire pour le moment cinétique. On peut montrer rigoureusement, grâce à ces deux équations, que l'orbite de l'électron reste circulaire pendant sa chute sur le noyau. Cette propriété est relativement évidente, puisque la force de réaction de rayonnement est opposée à la vitesse et de module constant.

Pour le mouvement Coulombien, on a, en utilisant le théorème du viriel:

$$\mathcal{E} = \frac{1}{2}\overline{U} = -\frac{q^2}{8\pi\epsilon_0}\frac{1}{r} \tag{1.52}$$

$$\frac{d\mathcal{E}}{dt} = \frac{q^2}{8\pi\epsilon_0 r^2}\frac{dr}{dt} \ , \tag{1.53}$$

pour une orbite circulaire. On a aussi, toujours pour une orbite circulaire:

$$\overline{\left(\frac{dU}{dr}\right)^2} = \frac{q^4}{(4\pi\epsilon_0)^2 r^4} \ . \tag{1.54}$$

On en déduit immédiatement que le rayon r de l'orbite doit décroître lentement en étant régi par l'équation:

$$\frac{dr}{dt} = -\frac{\tau}{m}\frac{q^2}{2\pi\epsilon_0 r^2} \ . \tag{1.55}$$

La résolution de cette équation est immédiate:

$$r^3(t) - r_0^3 = -9(c\tau)^3\frac{t}{\tau} \ . \tag{1.56}$$

Le cube du rayon de l'orbite décroît donc linéairement avec le temps. De manière évidente, le rayon s'annule en un temps fini: la chute de l'électron sur le noyau prend un temps fini. Au cours de cette chute, la fréquence orbitale et donc la fréquence du rayonnement émis croissent indéfiniment. Il est évident que nos approximations (faible rayonnement, mouvement non relativiste, approximation séculaire) tomberont avant la fin du mouvement. On peut cependant estimer l'ordre de grandeur de la durée de vie de l'atome d'hydrogène dans son état fondamental en prenant pour rayon initial le rayon de Bohr a_0. On trouve alors une durée de vie:

$$T = \frac{1}{9}\tau\left(\frac{a_0}{c\tau}\right)^3 \ , \tag{1.57}$$

qui est de l'ordre de τ multiplié par le cube du rapport du rayon de Bohr au rayon classique de l'électron. Numériquement, T vaut environ 14 ps. Un ordre de grandeur aussi faible montre bien les limites d'une approche classique à la structure atomique: au bout de quelques picosecondes tous les atomes auraient du disparaître dans un flash de rayonnement ultraviolet...

Notons que l'on peut, de manière plus utile, estimer à partir de ce calcul la durée de vie de l'orbite de Bohr de nombre quantique principal n. L'émission fait passer l'atome du niveau n au niveau $n-1$. Le rayon de l'orbite n étant $n^2 a_0$, la variation du cube du rayon dans cette transition est $6a_0^3 n^5$ (on

suppose que n est grand et on remplace la différence finie par une dérivée). La durée de vie de l'état n doit alors être de l'ordre de:

$$\tau \frac{2}{3}\left(\frac{a_0}{c\tau}\right)^3 n^5 \simeq 6n^5 T \ . \tag{1.58}$$

Elle croît très rapidement avec n. Pour $n = 50$, par exemple, on trouve 26 ms. Ce résultat est extrêmement proche de la durée de vie du niveau de Rydberg circulaire $n = 50$. Ce niveau, de nombres quantiques orbital et magnétique maximums, correspond en effet de la manière la plus proche possible à l'orbite circulaire de Bohr. Comme tous les nombres quantiques mis en jeu sont élevés, le principe de correspondance nous indique en effet que le calcul classique et le calcul quantique doivent être en très bon accord.

1.4 Rayonnement du dipôle

Nous allons consacrer cette section à l'étude, au moyen des potentiels de Liénard–Wiechert, d'un cas particulier important: le rayonnement d'une charge oscillant de manière harmonique au voisinage de l'origine. Au prix d'une approximation simple, valable quand l'extension du mouvement est très petite par rapport à la longueur d'onde rayonnée, nous verrons qu'il est possible de calculer explicitement les champs. Ce rayonnement dipolaire se retrouve dans de nombreux domaines. Comme nous le verrons à la fin de cette partie, le rayonnement des sources atomiques entre en général dans le cadre de cette approximation. Tous les problèmes d'émission ou de diffusion de rayonnement par des atomes ou des assemblées d'atomes se ramènent donc à un problème de rayonnement dipolaire. De plus, comme nous le verrons dans le prochain chapitre, ce rayonnement est le premier terme d'un développement (dit "multipolaire") valable pour une répartition quelconque de courants. Nous consacrerons la première partie de ce paragraphe à adapter simplement les résultats des paragraphes précédents à ce cas simple: calculs des potentiels, des champs rayonnés, du diagramme de rayonnement et de la puissance totale rayonnée. Dans la deuxième partie, nous exploiterons la très grande simplicité de cette source pour calculer les champs exacts. Nous conserverons donc tous les termes en $1/R^n$ et montrerons comment on passe des champs statiques aux champs rayonnés quand on s'éloigne de la source.

1.4.1 Champs rayonnés

Nous considérons donc une charge ponctuelle animée d'un mouvement harmonique *non relativiste* autour de l'origine O. Nous choisirons un mouvement linéaire le long de l'axe Oz. Un mouvement harmonique plus complexe peut en effet toujours être décrit comme la superposition de trois mouvements harmoniques selon les trois directions de l'espace. Dans le cadre des approximations que nous ferons, le champ correspondant est simplement la somme des champs créés indépendamment par ces trois mouvements.

La position de la particule à l'instant t' est donc:

$$\mathbf{r}_0(t') = z_0 \mathbf{u}_z e^{-i\omega t'} \tag{1.59}$$

et sa vitesse

$$\mathbf{v}(t') = -i z_0 \omega \mathbf{u}_z e^{-i\omega t'} \ . \tag{1.60}$$

Nous supposerons la vitesse très petite devant c et ne garderons que les termes dominants dans un développement en puissances de $\mathbf{v}/c$. En particulier, le facteur cinématique $1/g'$ qui est si important pour les particules relativistes sera simplement pris égal à 1. On peut s'interroger sur l'appellation dipôle pour une charge oscillante unique. Un vrai dipôle oscillant serait constitué de cette charge oscillante et d'une charge opposée fixe à l'origine. La charge fixe ne crée qu'un champ électrostatique. Comme nous ne nous intéressons qu'au champ rayonné à grande distance à la fréquence ω, cette contribution électrostatique ne joue aucun rôle. Seule la charge mobile importe pour le rayonnement.

Notons également que nous devrons, en toute rigueur, appeler cette source un dipôle électrique. Nous verrons en effet dans le prochain chapitre qu'il existe un autre type de source dipolaire, le dipôle magnétique, essentiellement équivalent à une petite boucle de courant oscillant.

L'approximation essentielle que nous ferons (approximation dipolaire) consiste à négliger l'extension spatiale du mouvement de la particule, z_0, par rapport à toutes les grandeurs caractéristiques du problème. Elle est tout d'abord négligeable par rapport à la distance d'observation et on écrira simplement $1/R = 1/r$ (r est le module de $\mathbf{r}$, distance de l'origine au point d'observation). Nous supposerons aussi que le temps retardé varie peu devant la période d'oscillation d'une extrémité de la trajectoire à l'autre. En d'autre termes, nous supposerons

$$\frac{z_0}{c} \ll \frac{1}{\omega} \tag{1.61}$$

ou encore

$$z_0 \ll \lambda \ , \tag{1.62}$$

où $\lambda = 2\pi c/\omega$ est la longueur d'onde, dans le vide, d'une onde plane de fréquence ω. Notons que cette condition s'écrit aussi $v_0 \ll c$. Nous pourrons alors écrire simplement:

$$t_0 = t - \frac{r}{c} \ . \tag{1.63}$$

t_0 est alors simplement retardé d'une quantité constante par rapport à l'instant d'observation. Nous écrivons ainsi le premier terme d'un développement de t_0 en puissances de r_0/λ. Notons que cette approximation dipolaire est, par exemple, tout à fait légitime pour le rayonnement atomique. Les longueurs d'onde caractéristiques sont en effet de l'ordre du micron alors que l'extension des mouvements électroniques, comparable au rayon de Bohr, est plutôt de l'ordre de l'Ångström.

Potentiels

Avec ces approximations, l'écriture des potentiels de Liénard ne pose aucune difficulté. Pour le potentiel vecteur, on a:

$$\mathbf{A} = \frac{\mu_0 j}{4\pi}\frac{e^{ikr}}{r}e^{-i\omega t}\frac{1}{1-\mathbf{n}\cdot\mathbf{v}/c}\mathbf{u}_z \ , \tag{1.64}$$

où l'on a utilisé l'amplitude j du courant créé par la particule:

$$je^{-i\omega t} = qve^{-i\omega t} = -iqz_0\omega e^{-i\omega t} \ , \tag{1.65}$$

Pour le potentiel scalaire:

$$V = \frac{q}{4\pi\epsilon_0}\frac{1}{r}\frac{1}{1-\mathbf{n}\cdot\mathbf{v}/c} \ . \tag{1.66}$$

Nous allons effectuer un développement de ces potentiels au premier ordre non trivial et non nul en v/c, en utilisant pour toutes les quantités retardées le seul instant t_0. On peut donc écrire:

$$\frac{1}{1-\mathbf{n}\cdot\mathbf{v}/c} \simeq 1 + \mathbf{n}\cdot\mathbf{v}/c \ . \tag{1.67}$$

Le premier ordre non nul pour le potentiel scalaire correspond au potentiel électrostatique $q/4\pi\epsilon_0 r$. Comme nous n'avons considéré qu'une charge ponctuelle, il est normal qu'on trouve effectivement à l'ordre le plus bas un potentiel électrostatique. Toutefois, il ne contribue pas à la propagation et au rayonnement et nous pouvons l'oublier (on pourrait aussi ajouter une charge $-q$ immobile à l'origine, ce qui constituerait une description plus réaliste d'un dipôle).

En remarquant que

$$\mathbf{n}\cdot\mathbf{v}/c = -i\frac{z_0\omega}{c}e^{i(kr-\omega t)}\mathbf{u}_z\cdot\mathbf{u}_r \ , \tag{1.68}$$

et que $\mathbf{u}_z \cdot \mathbf{u}_r = \cos\theta$ en coordonnées sphériques, on peut écrire les potentiels au premier ordre significatif:

$$\mathbf{A} = \frac{\mu_0 j}{4\pi}\frac{e^{ikr}}{r}e^{-i\omega t}\mathbf{u}_z \tag{1.69}$$

$$V = -i\frac{d}{4\pi\epsilon_0}\frac{\omega}{c}\frac{e^{i(kr-\omega t)}}{r}\cos\theta \ . \tag{1.70}$$

Dans l'expression de V, nous avons fait intervenir l'amplitude du dipôle associé à la charge:

$$d = ij/\omega = qz_0 \ . \tag{1.71}$$

On peut vérifier par un calcul élémentaire que V et $\mathbf{A}$ vérifient bien la jauge de Lorentz (les potentiels de Liénard Wiechert complets y obéissent. Ce doit donc être vrai à tous les ordres dans le développement en v/c).

Notons que nous aurions pu pousser le développement un terme plus loin en puissances de v/c. En fait, sauf pour vérifier la consistance de la jauge, il n'est pas nécessaire de calculer le potentiel scalaire au second ordre. Comme nous le verrons dans un moment, le champ magnétique se calcule directement à partir du potentiel vecteur et le champ électrique s'en déduit en utilisant les équations de Maxwell ou, plus simplement, la structure locale d'onde plane à très grande distance. Le terme d'ordre 2 dans le potentiel vecteur s'écrit, après quelques transformations:

$$\mathbf{A}^{(2)} = -\frac{\mu_0}{4\pi}\frac{qz_0^2}{2}\frac{4\omega^2}{2c}\frac{e^{2i(kr-\omega t)}}{r}\cos\theta\mathbf{u}_z \ . \tag{1.72}$$

Il s'agit d'un terme oscillant à la fréquence 2ω (c'est un terme en v^2), possédant une structure géométrique plus complexe que le potentiel vecteur du dipôle. Nous examinerons plus tard, par la technique systématique des développements multipolaires, le champ créé par une charge oscillante à cet ordre. Nous verrons qu'il existe effectivement un champ quadripolaire électrique à cette fréquence et nous retrouverons, par une autre méthode, l'expression ci-dessus.

Champs rayonnés

Les champs rayonnés s'expriment également sans difficultés. On pourrait calculer le champ électrique à partir de $\mathbf{A}$ et de V. Le calcul ne présente aucune difficulté. Toutefois, si ne nous intéressons qu'aux champs rayonnés à grande distance, il est beaucoup plus économique de revenir directement aux champ de Liénard. A l'ordre d'approximation où nous travaillons:

$$\mathbf{E} = \frac{q}{4\pi\epsilon_0 c^2}\frac{1}{r}\mathbf{n}\times(\mathbf{n}\times\mathbf{a}) \ , \tag{1.73}$$

où l'accélération $\mathbf{a}$ s'écrit simplement:

$$\mathbf{a} = -z_0\omega^2 e^{-i\omega t}\mathbf{u}_z \ . \tag{1.74}$$

Le champ électrique s'exprime plus naturellement en termes du dipôle électrique. On a

$$\mathbf{E} = -\frac{d}{4\pi\epsilon_0}k^2\frac{e^{ikr}}{r}e^{-i\omega t}\mathbf{n}\times(\mathbf{n}\times\mathbf{u}_z) \tag{1.75}$$

En utilisant les vecteurs unitaires du trièdre local en coordonnées sphériques, on trouve rapidement que:

$$\mathbf{n}\times\mathbf{u}_z = -\sin\theta\mathbf{u}_\phi \tag{1.76}$$

$$\mathbf{n}\times(\mathbf{n}\times\mathbf{u}_z) = \sin\theta\mathbf{u}_\theta \ , \tag{1.77}$$

et finalement

$$\mathbf{E} = -\frac{d}{4\pi\epsilon_0}\frac{\omega^2}{c^2}\frac{e^{ikr}}{r}e^{-i\omega t}\sin\theta\mathbf{u}_\theta \ . \tag{1.78}$$

L'onde rayonnée ayant toujours localement la structure d'une onde plane de vecteur d'onde $\mathbf{k} = k\mathbf{n}$, on en déduit

$$\mathbf{B} = -\frac{d}{4\pi\epsilon_0}\frac{\omega^2}{c^3}\frac{e^{ikr}}{r}e^{-i\omega t}\sin\theta\mathbf{u}_\phi \ . \tag{1.79}$$

La géométrie des champs est relativement intuitive. La répartition de courant produite par la charge est en effet à symétrie cylindrique autour de l'axe Oz. Le champ magnétique respecte cette symétrie (les lignes de champ sont des cercles d'axe Oz). Quand au champ électrique, il est dans le plan défini par le dipôle et le point d'observation. Notons enfin que les champs sont nuls dans la direction du dipôle, Oz. Nous discuterons dans un moment quelques conséquences physiques de cet effet non–relativiste (le rayonnement n'est pas du tout concentré dans la direction de la vitesse de la particule).

Vecteur de Poynting

A partir des champs, nous pouvons facilement calculer le vecteur de Poynting:

$$\mathbf{\Pi} = \frac{\mathbf{E}\times\mathbf{B}}{\mu_0} \ . \tag{1.80}$$

Il faut prendre garde aux notations complexes quand on manipule des quantités énergétiques, quadratiques dans les champs. Une substitution directe et maladroite des expressions précédentes pourrait faire croire que $\mathbf{\Pi}$ est une fonction purement oscillante à 2ω. Il faut, en fait, soit revenir aux parties réelles des champs pour exprimer la valeur instantanée de $\mathbf{\Pi}$, soit ne s'intéresser qu'à la moyenne temporelle $\overline{\mathbf{\Pi}}$ de $\mathbf{\Pi}$ sur une période optique (ou sur un intervalle de temps long par rapport à cette période). En effet, cette moyenne temporelle s'écrit simplement en termes des amplitudes complexes des champs. En posant:

$$\mathbf{E} = \mathbf{E}_0 e^{-i\omega t} \tag{1.81}$$

$$\mathbf{B} = \mathbf{B}_0 e^{-i\omega t} \ , \tag{1.82}$$

les champs réels s'écrivent

$$\mathbf{E} = \frac{1}{2}(\mathbf{E}_0 e^{-i\omega t} + \mathbf{E}_0^* e^{i\omega t}) \tag{1.83}$$

$$\mathbf{B} = \frac{1}{2}(\mathbf{B}_0 e^{-i\omega t} + \mathbf{B}_0^* e^{i\omega t}) \ . \tag{1.84}$$

($\mathbf{E}_0^*$ désignant le complexe conjugué de $\mathbf{E}_0$). Le vecteur de Poynting contient alors des termes oscillants à 2ω dont la moyenne temporelle est évidemment nulle et des termes constants, qui contribuent seuls à cette moyenne. On a finalement:

$$\overline{\mathbf{\Pi}} = \frac{1}{4\mu_0}\left[\mathbf{E}_0\times\mathbf{B}_0^* + \mathbf{E}_0^*\times\mathbf{B}_0\right] = \frac{1}{2\mu_0}\mathrm{Re}\left(\mathbf{E}_0\times\mathbf{B}_0^*\right) \ . \tag{1.85}$$

On notera que ce type d'expression peut être utilisé pour toute quantité quadratique (densités d'énergie...).

En appliquant cette formule nous trouvons donc:

$$\overline{\mathbf{\Pi}} = \frac{d^2}{32\pi^2\epsilon_0}\frac{\omega^4}{c^3}\frac{\sin^2\theta}{r^2}\mathbf{u}_r \ . \tag{1.86}$$

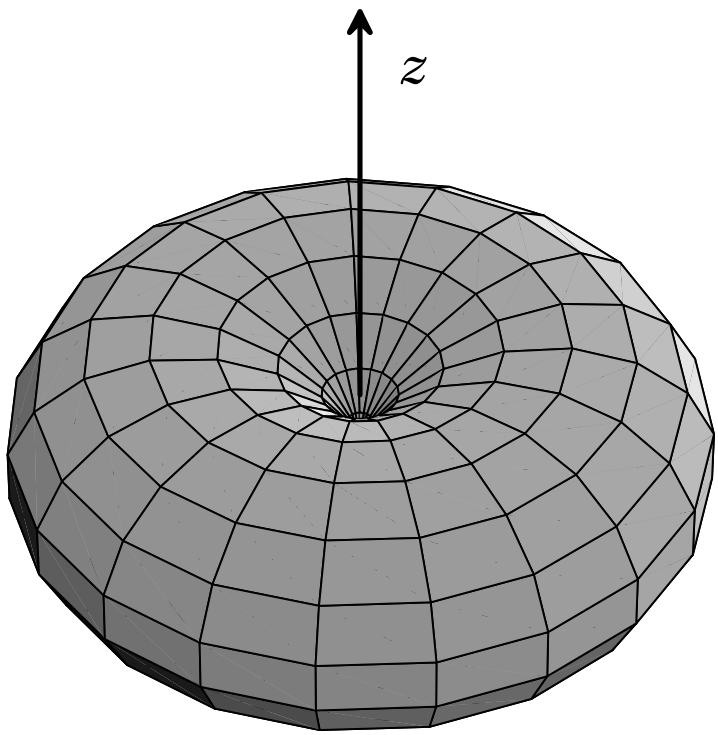

Figure 1.3: Diagramme de rayonnement d'un dipôle électrique aligné avec l'axe Oz. On note que le rayonnement est nul sur l'axe du dipôle.

Comme prévu, $\overline{\mathbf{\Pi}}$ est colinéaire à $\mathbf{u}_r$ et de même sens. Il décrit bien un flux d'énergie du dipôle vers l'infini. On peut visualiser la répartition spatiale du flux énergétique en traçant le diagramme de rayonnement. Dans la direction définie par les angles θ et ϕ des coordonnées sphériques, on porte une longueur proportionnelle à $\overline{\mathbf{\Pi}}$. Le diagramme ainsi obtenu est de révolution autour de Oz. Il est représenté sur la figure 1.3. La section de cette surface par un plan passant par Oz est constituée de deux lobes symétriques d'équation polaire $\rho = K \sin^2 \theta$. La propriété essentielle de ce diagramme est que la puissance rayonnée est nulle dans la direction du mouvement, maximale dans une direction perpendiculaire.

De nombreux phénomènes physiques sont associés à ce diagramme de rayonnement. Une antenne dipolaire électrique (un fil rectiligne parcouru par un courant oscillant — nous montrerons dans le prochain chapitre l'équivalence avec le système traité ici) ne rayonne que dans un plan perpendiculaire à son axe. Elle doit donc être orientée verticalement pour une communication terrestre. De manière réciproque, une antenne linéaire ne capte correctement le rayonnement que si elle est perpendiculaire à la direction de propagation et parallèle à la polarisation (une expérience quotidienne avec les antennes de télévision, réseau d'antennes linéaires). Nous verrons qu'une petite particule ou un atome éclairé par une onde plane possède un dipôle induit parallèle à la direction de polarisation. Il diffuse donc préférentiellement le rayonnement dans une direction perpendiculaire à sa polarisation. Ce phénomène est à l'origine de la polarisation de la lumière diffusée par l'atmosphère. Les ondes incidentes de polarisation perpendiculaire au plan défini par la direction d'observation et la direction du soleil contribuent principalement à la lumière diffusée. La lumière observée dans une direction est donc partiellement polarisée. Quand la lumière subit des diffusion multiples (par temps de brouillard par exemple), cet effet de polarisation est brouillé et la lumière diffusée a une polarisation naturelle. On peut observer facilement la direction préférentielle de diffusion, dans une expérience de cours, en observant la diffusion d'une lumière polarisée par une solution saturée de glucose. Comme elle possède un important pouvoir rotatoire, la polarisation linéaire de l'onde incidente tourne, avec une période

spatiale de quelques centimètres. On observe alors, à angle droit avec la direction incidente, que la diffusion est pratiquement nulle quand la polarisation pointe vers l'observateur, maximale dans le cas contraire.

Puissance totale rayonnée

On obtient la moyenne temporelle de la puissance rayonnée par le dipôle, $\mathcal{P}$, en intégrant le vecteur de Poynting moyen sur une sphère de rayon r. $\overline{\mathbf{\Pi}}$ étant en $1/r^2$, la dépendance en r disparaît et la puissance rayonnée est indépendante du rayon de cette sphère. L'intégrale sur l'angle ϕ est triviale et on obtient immédiatement:

$$\mathcal{P} = \frac{d^2}{32\pi^2\epsilon_0}\frac{\omega^4}{c^3}2\pi\int_0^\pi \sin^3\theta\, d\theta \tag{1.87}$$

On vérifiera que

$$\int_0^\pi \sin^3\theta\, d\theta = \frac{4}{3}\ . \tag{1.88}$$

On a donc enfin:

$$\mathcal{P} = \frac{d^2}{12\pi\epsilon_0}\frac{\omega^4}{c^3}\ . \tag{1.89}$$

Cette formule nous sera très utile dans la suite de ce cours. De manière naturelle, la puissance rayonnée est proportionnelle au carré de l'amplitude du dipôle. Elle est également proportionnelle à la quatrième puissance de la fréquence. A dipôle égal, les hautes fréquences sont rayonnées de façon beaucoup plus efficace que les basses. On connaît bien ce phénomène en acoustique: les hauts–parleurs sont d'autant plus petits, pour une puissance restituée égale, que leur fréquence optimale d'utilisation est élevée. Dans le domaine optique, ce comportement est responsable de la couleur bleue de la lumière diffusée si la lumière incidente est blanche. C'est par exemple ce qui explique le bleu du ciel. Les molécules d'air ont des fréquences de résonance très hautes par rapport aux fréquences visibles. Elles répondent donc de la même manière à toutes les fréquences visibles, par des dipôles comparables. En revanche, les fréquences les plus élevées dominent largement dans le spectre de la lumière diffusée. Par un simple effet de complémentarité, le rouge domine dans le spectre de la lumière transmise. Voici pourquoi le ciel est bleu [2] et les couchers de soleil rouges.

La puissance rayonnée doit provenir d'une source d'énergie. On pourrait envisager que la particule soit mise en mouvement par des sources externes, éventuellement non électromagnétiques. L'énergie rayonnée peut aussi provenir de l'énergie mécanique de la particule (élastiquement liée à un centre de force). Dans ce cas, cette énergie mécanique décroît au cours du temps sous l'action de la force de réaction de rayonnement. Cette force, proportionnelle à la dérivée troisième de la position est, pour un mouvement pratiquement sinusoïdal, proportionnelle au cube de la fréquence. La puissance de cette force varie donc comme la quatrième puissance de la fréquence, ce qui est bien ce que nous venons de déterminer par un autre moyen.

1.4.2 Champs à une distance arbitraire

Le problème que nous nous sommes posé ici est suffisamment simple pour que nous puissions entreprendre de calculer exactement le champ produit à une distance arbitraire. Nous pourrons ainsi calculer le champ proche du dipôle et montrer comment on passe d'une solution essentiellement électrostatique à courte distance aux champ rayonnés du paragraphe précédent. Nous partirons du potentiel vecteur décrit par l'équation (1.69). La seule approximation effectuée à ce niveau est en effet l'approximation dipolaire négligeant l'extension du mouvement par rapport à la longueur d'onde rayonnée. En revanche

[2] L'explication complète du bleu du ciel est un peu plus complexe. Si l'atmosphère était parfaitement homogène, elle se comporterait comme un milieu d'indice constant et il n'y aurait pas de diffusion. Le ciel serait noir (et triste). Les fluctuations thermodynamiques de densité sont responsables d'inhomogénéités d'indice à l'échelle de la longueur d'onde incidente et de l'existence d'une diffusion.

les champs de l'équation (1.78,1.79) ne sont que les premiers termes d'un développement en puissances de $1/r$ des vrais champs. Nous allons rétablir intégralement ce développement. Nous commencerons par établir l'expression du champ magnétique, rotationnel du potentiel vecteur. Nous en déduirons ensuite le champ électrique, directement à partir de l'équation de Maxwell–Ampère. Nous éviterons ainsi d'avoir à expliciter le potentiel scalaire, qui est important dans la zone des champs proches.

Champ Magnétique

Nous partons de

$$\mathbf{B} = \boldsymbol{\nabla} \times \mathbf{A} = \frac{\mu_0 j}{4\pi} e^{-i\omega t} \boldsymbol{\nabla} \times \frac{e^{ikr}}{r} \mathbf{u}_z \ . \tag{1.90}$$

Une formule standard d'analyse vectorielle donne:

$$\boldsymbol{\nabla} \times \frac{e^{ikr}}{r} \mathbf{u}_z = \boldsymbol{\nabla} \left(\frac{e^{ikr}}{r} \right) \times \mathbf{u}_z \ , \tag{1.91}$$

le rotationnel de $\mathbf{u}_z$ étant manifestement nul. Le gradient étant celui d'une fonction à symétrie sphérique, il s'exprime simplement:

$$\boldsymbol{\nabla} \frac{e^{ikr}}{r} = \mathbf{u}_r e^{ikr} \left[\frac{ik}{r} - \frac{1}{r^2} \right] \ . \tag{1.92}$$

En remarquant que $\mathbf{u}_r \times \mathbf{u}_z = -\sin\theta \mathbf{u}_\phi$, on a finalement le champ magnétique sous la forme:

$$\mathbf{B} = -\frac{\mu_0 j}{4\pi} e^{-i\omega t} e^{ikr} \left[\frac{ik}{r} - \frac{1}{r^2} \right] \sin\theta \mathbf{u}_\phi \ . \tag{1.93}$$

Notons que la partie en $1/r$ de ce champ, dominante à grande distance, coïncide, à des réécritures près, avec (1.79). Notons également que le terme de champ proche, en $1/r^2$ ici, n'affecte pas les propriétés de symétrie du champ dont les lignes restent des cercles d'axe Oz.

Champ Electrique

Plutôt que d'écrire l'expression du potentiel scalaire, par exemple par la jauge de Lorentz, pour la dériver ensuite, nous allons utiliser l'équation de Maxwell–Ampère qui s'écrit, pour des champs harmoniques:

$$\mathbf{E} = \frac{ic^2}{\omega} \boldsymbol{\nabla} \times \mathbf{B} \ . \tag{1.94}$$

Il est plus naturel d'exprimer le champ électrique en termes de l'amplitude du dipôle $d = qz_0$. Après une simple réécriture, on a:

$$\mathbf{E} = \frac{d}{4\pi\epsilon_0} e^{-i\omega t} \boldsymbol{\nabla} \times \boldsymbol{\nabla} \times \left(\frac{e^{ikr}}{r} \mathbf{u}_z \right) \ . \tag{1.95}$$

Le double rotationnel se développe en:

$$\boldsymbol{\nabla} \times \boldsymbol{\nabla} \times \frac{e^{ikr}}{r} \mathbf{u}_z = \boldsymbol{\nabla} \left(\boldsymbol{\nabla} \cdot \frac{e^{ikr}}{r} \mathbf{u}_z \right) - \triangle \frac{e^{ikr}}{r} \mathbf{u}_z \ . \tag{1.96}$$

Le laplacien du second membre est particulièrement simple. $\mathbf{u}_z$ étant constant, il se ramène à:

$$\triangle \frac{e^{ikr}}{r} \mathbf{u}_z = \mathbf{u}_z \triangle \frac{e^{ikr}}{r} = \frac{1}{r} \frac{d^2}{dr^2} r \left(\frac{e^{ikr}}{r} \right) \mathbf{u}_z = -\frac{k^2}{r} e^{ikr} \mathbf{u}_z \ . \tag{1.97}$$

En utilisant enfin la décomposition de $\mathbf{u}_z$ sur le trièdre local ($\mathbf{u}_z = -\sin\theta\mathbf{u}_\theta + \cos\theta\mathbf{u}_r$), on a:

$$\triangle\frac{e^{ikr}}{r}\mathbf{u}_z = +\sin\theta\frac{k^2}{r}e^{ikr}\mathbf{u}_\theta - \cos\theta\frac{k^2}{r}e^{ikr}\mathbf{u}_r \tag{1.98}$$

Revenons maintenant au premier terme du membre de gauche de (1.96). $\mathbf{u}_z$ étant constant, il sort de la divergence et nous pouvons écrire:

$$\boldsymbol{\nabla}\left(\boldsymbol{\nabla}\cdot\frac{e^{ikr}}{r}\mathbf{u}_z\right) = \boldsymbol{\nabla}\left(\boldsymbol{\nabla}\frac{e^{ikr}}{r}\cdot\mathbf{u}_z\right) \ . \tag{1.99}$$

Le gradient d'un produit scalaire peut aussi se développer par une formule d'analyse vectorielle standard. En utilisant encore une fois le fait que $\mathbf{u}_z$ est une constante, et que ses dérivées sont nulles, ainsi que le fait que le rotationnel d'un gradient est identiquement nul, on a:

$$\boldsymbol{\nabla}\left(\boldsymbol{\nabla}\cdot\frac{e^{ikr}}{r}\mathbf{u}_z\right) = (\mathbf{u}_z\cdot\boldsymbol{\nabla})\boldsymbol{\nabla}\frac{e^{ikr}}{r} \ . \tag{1.100}$$

avec $(\mathbf{u}_z\cdot\boldsymbol{\nabla}) = \cos\theta\partial/\partial r - (\sin\theta/r)\partial/\partial\theta$, et en exprimant le gradient on obtient enfin:

$$\boldsymbol{\nabla}\left(\boldsymbol{\nabla}\cdot\frac{e^{ikr}}{r}\mathbf{u}_z\right) = \cos\theta\mathbf{u}_r e^{ikr}\left[-\frac{k^2}{r} - 2\frac{ik}{r^2} + \frac{2}{r^3}\right] - \sin\theta\mathbf{u}_\theta e^{ikr}\left[\frac{ik}{r^2} - \frac{1}{r^3}\right] \ . \tag{1.101}$$

En regroupant finalement ce terme avec celui provenant du laplacien on a le champ à toute distance sous la forme:

$$\mathbf{E} = \frac{d}{4\pi\epsilon_0}e^{-i\omega t}e^{ikr}\left\{-2\mathbf{u}_r\cos\theta\left[\frac{ik}{r^2} - \frac{1}{r^3}\right] - \mathbf{u}_\theta\sin\theta\left[\frac{k^2}{r} + \frac{ik}{r^2} - \frac{1}{r^3}\right]\right\} \ . \tag{1.102}$$

Nous avons donc bien exprimé le champ électrique comme un développement en puissances successives de kr ou de r/λ. Le terme en $1/r$ dominant à grande distance coïncide, là encore, avec le champ rayonné donné par (1.78). La symétrie du champ électrique n'est pas modifiée par les termes de champ proche et $\mathbf{E}$ reste dans le plan défini par la direction du dipôle et la direction d'observation. Notons enfin que la partie du champ en $1/r^2$, importante dans la région intermédiaire, est en quadrature avec les champs en $1/r$ et $1/r^3$. Cette propriété a une certaine importance dans les raisonnements énergétiques.

L'expression de $\mathbf{E}$ en termes des vecteurs de base du trièdre local est sans doute la plus commode. Notons pour mémoire qu'on peut donner une expression complètement vectorielle de $\mathbf{E}$ sous la forme:

$$\mathbf{E} = \frac{d}{4\pi\epsilon_0}e^{-i\omega t}e^{ikr}\left\{\frac{k^2}{r}(\mathbf{u}_r\times\mathbf{u}_z)\times\mathbf{u}_r - \left(\frac{ik}{r^2} - \frac{1}{r^3}\right)(3\mathbf{u}_r\times(\mathbf{u}_r\times\mathbf{u}_z) + 2\mathbf{u}_z)\right\} \ . \tag{1.103}$$

Cette forme a le mérite essentiel de bien faire ressortir la contribution rayonnée, mise ici sous la forme entièrement vectorielle résultant directement des champs de Liénard.

Discussion

Trois régions importantes de l'espace apparaissent dans l'équation (1.102). La région des champ rayonnés correspond à $kr \gg 1$ et nous l'avons déjà étudiée en détails. La région $kr \ll 1$ est dite région du champ proche. Le champ électrique y est dominé par des termes en $1/r^3$. Il peut s'écrire:

$$\mathbf{E} = \frac{d}{4\pi\epsilon_0}e^{-i\omega t}\left[\frac{2\cos\theta}{r^3}\mathbf{u}_r + \frac{\sin\theta}{r^3}\mathbf{u}_\theta\right] \ . \tag{1.104}$$

On reconnaît, à un facteur oscillant $e^{-i\omega t}$ près, l'expression du champ électrique du dipôle électrostatique. Nous sommes effectivement à une distance courte devant la longueur d'onde rayonnée: la propagation ne joue pas de rôle et le champ électrique suit immédiatement sa source.

Ce champ “électrostatique” proche joue un rôle important dans de nombreux domaines. Il est par exemple responsable des forces de “van der Waals” entre molécules d'un gaz, responsables en partie des écarts à la loi des gaz parfaits. Les molécules sont en effet de petits dipôles électrostatiques fluctuant à des fréquences optiques. Les distances moyennes entre particules dans un gaz de densité normale étant très inférieures aux longueurs d'onde optiques, l'interaction entre ces petits dipôles peut se calculer en utilisant ces formules de champ proche. Il serait hors de propos de donner ici une interprétation plus détaillée des forces de van der Waals, qui ne peuvent se comprendre correctement que dans un cadre quantique (on notera par exemple que l'image d'un dipôle fluctuant ne peut tenir pour une molécule dans son état fondamental: ce dipôle devrait rayonner, ce qui rendrait l'état fondamental instable).

La troisième région de l'espace est celle des champs “intermédiaires”, $kr \simeq 1$. Tous les termes du champ électrique sont alors d'importance comparable. La complexité du champ est telle que nous ne décrirons pas plus en détail ses propriétés.

Chapitre 2

Développement multipolaire du champ rayonné

Nous nous intéresserons dans ce chapitre à un type de sources qui généralise le dipôle introduit au chapitre précédent. Nous considérerons une répartition de courants quelconques, oscillant de façon harmonique, localisés au voisinage de l'origine. Les courants seront imposés, comme le mouvement des charges dans le chapitre précédent. Le problème ainsi posé est bien sûr trop général. Nous chercherons donc seulement à déterminer les champs rayonnés à une distance très grande par rapport à la longueur d'onde et à l'extension spatiale de la répartition de courants. Nous ne nous restreindrons pas en revanche à des répartitions de courant localisées sur une étendue faible par rapport à la longueur d'onde. Nous procéderons à un développement en puissances de l'extension de la source par rapport à λ (longueur d'onde rayonnée). Nous identifierons physiquement les différents termes de ce développement et constaterons que le terme dominant n'est autre que le champ du dipôle électrique. Nous préciserons à chaque fois les caractéristiques du rayonnement et, en particulier, le diagramme d'émission. Nous conclurons ce chapitre en présentant quelques applications de ces développements multipolaires. Nous reviendrons en particulier sur le rayonnement d'une charge animée d'un mouvement sinusoïdal non relativiste. Nous pourrons calculer les corrections au rayonnement dipolaire dues à l'extension finie de la trajectoire. Nous nous pencherons ensuite sur le problème, plus important en pratique, des antennes. Nous discuterons deux types d'antennes couramment répandues. Nous dégagerons en particulier la notion importante d'impédance de rayonnement.

2.1 Développement multipolaire du potentiel vecteur

2.1.1 Notations

La position du problème et les notations sont résumées sur la figure 2.1. Nous cherchons le champ rayonné en un point $\mathbf{r}$ à l'instant t par une répartition de courant située au voisinage de l'origine, dans une région d'extension caractéristique r_0 (on prendra garde à quelques changements de notation par rapport au paragraphe précédent). Le courant au point source $\mathbf{r}'$ à l'instant t' est $\mathbf{j}(\mathbf{r}', t') = \mathbf{j}_0(\mathbf{r}')\exp(-i\omega t')$. $\mathbf{j}_0$ désigne donc l'amplitude complexe de la densité de courant. Notons encore une fois que, si nous nous limitons ici à des courants harmoniques, le cas général peut être traité en utilisant transformation de Fourier et principe de superposition.

La répartition de courants est nécessairement, par la conservation de la charge, accompagnée d'une densité de charges, elle aussi oscillante à la fréquence ω: $\rho(\mathbf{r}', t') = \rho_0(\mathbf{r}')\exp(-i\omega t')$. En écrivant l'équation de continuité:

$$\boldsymbol{\nabla} \cdot \mathbf{j} + \frac{\partial \rho}{\partial t'} = 0 \tag{2.1}$$

où les dérivations spatiales s'entendent par rapport à $\mathbf{r}'$ et en y portant la dépendance harmonique en

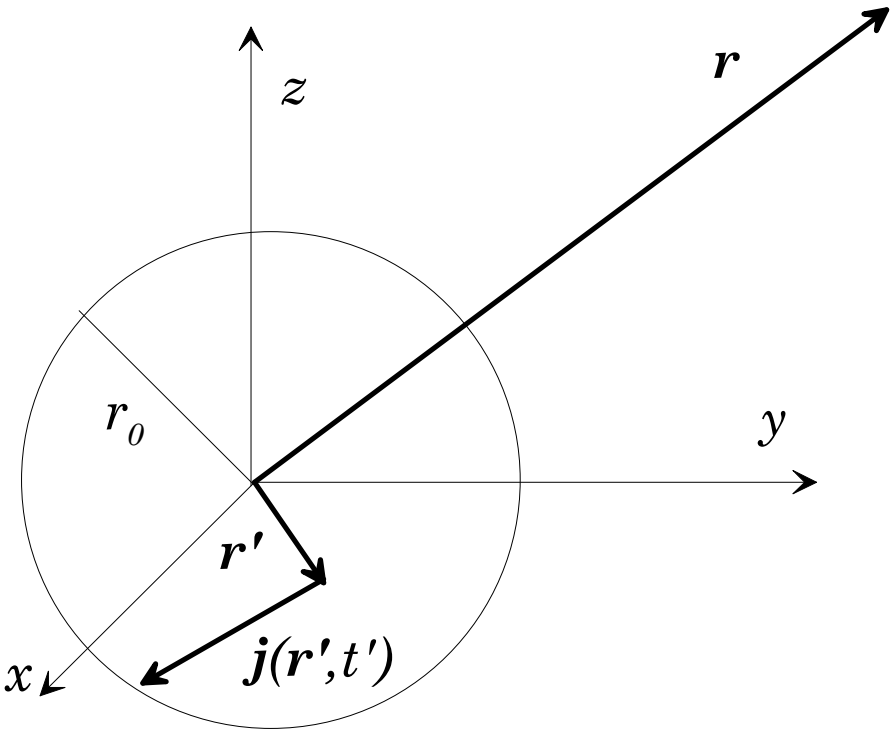

Figure 2.1: Rayonnement multipolaire: notations

temps, on trouve:

$$i\omega\rho_0 = \boldsymbol{\nabla}\cdot\mathbf{j}_0 \ . \tag{2.2}$$

Nous utiliserons donc indifféremment la densité de courant et la densité de charge pour caractériser la source.

Il y a dans ce problème trois échelles naturelles de longueur:

- La distance r entre la répartition de courant et le point d'observation.
- La longueur d'onde $\lambda = 2\pi c/\omega$ du rayonnement
- L'extension spatiale r_0 de la répartition de courant.

Nous nous placerons uniquement dans le cas où la distance d'observation est très grande, à la fois par rapport à la longueur d'onde et par rapport à l'extension de la source:

$$r \gg \lambda, r_0 \ . \tag{2.3}$$

En fait, nous calculerons le potentiel comme un développement limité en puissances de r/λ et nous ne garderons que les termes dominants, en $1/r$. Nous suivons en cela la démarche du chapitre précédent. Nous ne supposerons pas en revanche, comme pour le dipôle, que $r_0 \ll \lambda$. Nous effectuerons plutôt un développement limité en puissances de r_0/λ, dont nous garderons tous les ordres [1] (le développement multipolaire proprement dit). Nous identifierons ensuite les contenus physiques de ces différents ordres. Pour des raison évidentes, nous nous cantonnerons aux termes des premier et second ordres pour cette analyse détaillée.

[1] Le problème du rayon de convergence d'un tel développement en série se pose bien sûr. En fait, ce développement est toujours convergent puisqu'il s'agit, comme nous le verrons, du simple développement d'une exponentielle

2.1.2 Potentiel vecteur

L'expression du potentiel vecteur s'obtient tout naturellement à partir de la solution en termes de potentiels retardés:

$$\begin{aligned}\mathbf{A}(\mathbf{r},t) &= \frac{\mu_0}{4\pi}\int \frac{\mathbf{j}_0(\mathbf{r}')e^{-i\omega(t-|\mathbf{r}-\mathbf{r}'|/c)}}{|\mathbf{r}-\mathbf{r}'|}\,d^3\mathbf{r}' \\ &= \frac{\mu_0}{4\pi}e^{-i\omega t}\int \frac{\mathbf{j}_0(\mathbf{r}')e^{ik|\mathbf{r}-\mathbf{r}'|}}{|\mathbf{r}-\mathbf{r}'|}\,d^3\mathbf{r}' \ , \end{aligned} \tag{2.4}$$

en posant encore $k = \omega/c$. Cette expression est, jusqu'ici, exacte pour une répartition de courant harmonique (rappelons que cette expression était, pour une dépendance temporelle quelconque, notre point de départ pour l'établissement des potentiels de Liénard).

Suivant notre programme, nous commençons par un développement limité en puissances de $1/r$ dont nous ne garderons que les termes dominants. La dépendance en $|\mathbf{r}-\mathbf{r}'|$ intervient dans le facteur d'amplitude (en $1/r$) et dans la phase due au temps retardé. Clairement, la dépendance en phase est beaucoup plus critique: il suffit que $|\mathbf{r}-\mathbf{r}'|$ varie de λ pour que le terme de phase change beaucoup. Si nous nous contentons de garder l'ordre 1 en r'/r dans la phase, il est cohérent de ne garder que l'ordre 0 pour l'amplitude et d'écrire $|\mathbf{r}-\mathbf{r}'| = r$ (Nous laisserons le soin au lecteur, à titre d'exercice, de vérifier la cohérence de ce développement). Le terme de phase peut s'écrire au premier ordre:

$$ik|\mathbf{r}-\mathbf{r}'| = ikr\sqrt{1-\frac{r'^2}{r^2}-2\frac{\mathbf{r}\cdot\mathbf{r}'}{r^2}} = ikr - ik\mathbf{u}_r\cdot\mathbf{r}' \ , \tag{2.5}$$

où $\mathbf{u}_r$ est le vecteur unitaire de la direction d'observation. En reportant ces deux développements dans l'expression du potentiel vecteur, on trouve:

$$\mathbf{A} = \frac{\mu_0}{4\pi}\frac{e^{i(kr-\omega t)}}{r}\int \mathbf{j}_0(\mathbf{r}')e^{-ik\mathbf{u}_r\cdot\mathbf{r}'}\,d^3\mathbf{r}' \ . \tag{2.6}$$

Comme nous pouvions nous y attendre, la dépendance en $\mathbf{r}$ du potentiel vecteur est essentiellement celle d'une onde sphérique. La complexité de la source est complètement contenue dans l'intégrale. Si cette intégrale ne porte que sur le point source $\mathbf{r}'$, le point d'observation y intervient par le vecteur $\mathbf{u}_r$.

Nous allons maintenant développer cette intégrale en puissances de l'extension de la source r_0 comparée à la longueur d'onde rayonnée λ. En raison de la densité de courant, l'intégrale ne porte que sur une région de l'espace d'extension r_0. Dans cette région finie, on peut développer l'exponentielle en:

$$e^{-ik\mathbf{u}_r\cdot\mathbf{r}'} = 1 - ik\mathbf{u}_r\cdot\mathbf{r}' + \frac{(-ik\mathbf{u}_r\cdot\mathbf{r}')^2}{2!} + \dots \ . \tag{2.7}$$

Notons que ce développement en série a un rayon de convergence infini. L'écrire ne présuppose en rien que l'extension de la source soit petite devant λ. En revanche, la convergence ne sera suffisamment rapide pour que le développement soit utile que si cette extension n'est pas trop grande. Nous écrirons donc enfin:

$$\mathbf{A} = \sum_{p=0}^{\infty}\mathbf{A}_p \ , \tag{2.8}$$

avec

$$\mathbf{A}_p = \frac{\mu_0}{4\pi}\frac{e^{i(kr-\omega t)}}{r}\frac{(-ik)^p}{p!}\int \mathbf{j}_0(\mathbf{r}')(\mathbf{u}_r\cdot\mathbf{r}')^p\,d^3\mathbf{r}' \ . \tag{2.9}$$

Nous appellerons développement multipolaire du potentiel vecteur cette expression. Nous allons en effet voir dans les paragraphes suivants, en examinant les termes successifs, qu'ils correspondent à des géométries de sources dipolaires, quadripolaires etc..

2.2 Termes multipolaires

Nous allons maintenant examiner les termes d'ordres successifs dans le développement (2.9). Si l'extension de la source n'est pas très grande devant λ, les premiers termes doivent dominer le potentiel. Pour des raisons évidentes de taille des calculs, nous nous limiterons en fait aux termes d'ordre zéro et un. On peut traiter de manière systématique les termes d'ordre arbitraire en les reliant aux propriétés des harmoniques sphériques. Cette méthode très puissante, bien adaptée à l'utilisation des propriétés de symétrie de la source, est exposée en détails dans le Jackson. L'exposer ici nous conduirait à une digression mathématique trop longue. De plus, l'essentiel des idées physiques est contenu dans la méthode que nous présentons. Pour chaque terme, nous calculerons séparément les champs électrique et magnétique rayonnés. Les théorèmes de superposition nous permettront de trouver les champs rayonnés par des sources quelconques. En revanche, les considérations énergétiques que nous pourrons faire pour des termes séparés, y compris l'établissement du diagramme de rayonnement, ne tiendront que si ces termes sont dominants (il n'y a pas de principe de superposition pour ces quantités quadratiques). Comme nous nous sommes placés dès l'abord dans le domaine des champs rayonnés, nous pourrons obtenir les champs à partir des potentiels en utilisant la structure locale d'onde plane de l'onde rayonnée. En fait, nous raisonnerons aussi souvent par analogie en montrant qu'il existe des liens très forts entre la géométrie des champs correspondant à différents termes.

2.2.1 Ordre 0: Dipôle électrique

Le terme d'ordre zéro du développement s'écrit simplement:

$$\mathbf{A}_0 = \frac{\mu_0}{4\pi}\frac{e^{i(kr-\omega t)}}{r}\int \mathbf{j}_0(\mathbf{r}')\,d^3\mathbf{r}' \ . \tag{2.10}$$

A cet ordre d'approximation, tout se passe comme si on avait un courant oscillant complètement localisé à l'origine, ayant comme valeur l'intégrale de volume du courant de la répartition initiale. Nous allons montrer, par quelques transformations algébriques simples de l'intégrale, qu'on retrouve en fait le potentiel vecteur produit par un dipôle électrique oscillant, modélisé, comme au chapitre précédent, par une charge oscillante.

Considérons donc la composante selon $\mathbf{u}_x$ de l'intégrale. Elle peut s'écrire:

$$I_x = \int \mathbf{u}_x \cdot \mathbf{j}_0(\mathbf{r}')\,d^3\mathbf{r}' = \int \boldsymbol{\nabla} x' \cdot \mathbf{j}_0(\mathbf{r}')\,d^3\mathbf{r}' \tag{2.11}$$

où le gradient s'entend par rapport à $\mathbf{r}'$. Une intégration par parties donne alors:

$$I_x = -\int x'(\boldsymbol{\nabla}\cdot\mathbf{j}_0)\,d^3\mathbf{r}' = -\int x'(i\omega\rho_0)\,d^3\mathbf{r}' \tag{2.12}$$

où nous avons cette fois utilisé l'équation de conservation de la charge pour faire intervenir la densité de charge. I_x apparaît donc comme la composante selon $\mathbf{u}_x$ d'un vecteur $-i\omega\mathbf{d}_0$ avec:

$$\mathbf{d}_0 = \int \mathbf{r}'\rho_0(\mathbf{r}')\,d^3\mathbf{r}' \ . \tag{2.13}$$

Pour une densité de charges statique, nous reconnaîtrions ici l'expression du moment dipolaire électrique. Pour notre répartition oscillant de façon harmonique, $\mathbf{d}_0$ est évidemment l'amplitude complexe du moment dipolaire électrique.

On a donc:

$$\mathbf{A}_0 = \frac{\mu_0}{4\pi}\frac{e^{i(kr-\omega t)}}{r}(-i\omega)\mathbf{d}_0 \tag{2.14}$$

On reconnaît ici l'expression du potentiel de la charge oscillante 1.69 (on prendra garde de remplacer dans l'expression présente le dipôle $\mathbf{d}_0$ par celui de la charge oscillante $qa\mathbf{u}_z$ pour comparer ces deux expressions).

Le terme d'ordre zéro du développement multipolaire correspond donc au dipôle électrique, pour lequel nous avons déjà calculé en détails les champs[2] et la puissance rayonnés. Nous ne discuterons pas à nouveau de ces termes. En général, une source possède un dipôle électrique non nul. Si l'extension est petite devant λ, ou même de l'ordre de λ comme nous le verrons dans les paragraphes suivants, ce terme domine. La grande majorité des sources se comportent comme des dipôles, ce qui justifie largement la place que nous avons accordée au traitement de ce problème. Ce n'est que pour des sources de grande extension ou pour celles qui, souvent pour des raisons de symétrie, ont un dipôle électrique strictement nul que les termes suivants, que nous allons discuter maintenant, jouent un rôle.

2.2.2 Ordre 1: Dipôle magnétique, Quadripôle électrique

Quand le dipôle électrique d'une distribution est nul, le terme d'ordre 1 domine. Le potentiel vecteur correspondant s'écrit:

$$\mathbf{A}_1 = \frac{\mu_0}{4\pi}\frac{e^{i(kr-\omega t)}}{r}(-ik)\int (\mathbf{u}_r\cdot\mathbf{r}')\mathbf{j}_0(\mathbf{r}')\,d^3\mathbf{r}' \ . \tag{2.15}$$

La transformation de cette intégrale en quelque chose de manipulable n'est pas aisée sous cette forme. Dans le produit d'un vecteur par un produit scalaire, on reconnaît un des termes du développement d'un double produit vectoriel. Nous pouvons effectivement transformer l'intégrande pour faire apparaître ce produit vectoriel, au prix de termes soustractifs supplémentaires. Nous écrirons pour cela:

$$\begin{aligned}(\mathbf{u}_r\cdot\mathbf{r}')\mathbf{j}_0 &= \frac{1}{2}\left[(\mathbf{u}_r\cdot\mathbf{r}')\mathbf{j}_0-(\mathbf{u}_r\cdot\mathbf{j}_0)\mathbf{r}'\right]\\ &\quad+\frac{1}{2}\left[(\mathbf{u}_r\cdot\mathbf{r}')\mathbf{j}_0+(\mathbf{u}_r\cdot\mathbf{j}_0)\mathbf{r}'\right] \ .\end{aligned} \tag{2.16}$$

La première ligne de ce développement, a priori très artificiel, est alors un double produit vectoriel. Nous allons voir, en discutant chacun des termes associés à chacune des lignes de (2.16), que cette séparation correspond à deux types de sources de propriétés physiques différentes, intervenant au même ordre dans le développement multipolaire: le dipôle magnétique et le quadripôle électrique. Nous écrirons donc le potentiel vecteur à l'ordre 1 comme:

$$\mathbf{A}_1 = \mathbf{A}_1^{dm} + \mathbf{A}_1^{qe} \ , \tag{2.17}$$

où $\mathbf{A}_1^{dm}$ correspond à la première ligne de (2.16) alors que $\mathbf{A}_1^{qe}$ correspond à la deuxième

Dipôle magnétique

On a donc:

$$\mathbf{A}_1^{dm} = \frac{\mu_0}{4\pi}\frac{e^{i(kr-\omega t)}}{r}(-ik)\int \frac{1}{2}\mathbf{u}_r\times(\mathbf{j}_0\times\mathbf{r}')\,d^3\mathbf{r}' \ . \tag{2.18}$$

Nous poserons donc (en remarquant que $\mathbf{u}_r$ sort de l'intégrale sur $\mathbf{r}'$):

$$\boldsymbol{\mathcal{M}}_0 = \frac{1}{2}\int \mathbf{r}'\times\mathbf{j}_0(\mathbf{r}')\,d^3\mathbf{r}' \ . \tag{2.19}$$

Pour une répartition de courants statiques, $\boldsymbol{\mathcal{M}}_0$ représenterait le dipôle magnétique de la répartition[3]. L'expression ci–dessus généralise simplement cette définition à l'amplitude complexe d'un dipôle

[2] On prendra garde que l'expression des champs à courte distance, que nous avions établie dans le chapitre précédent, n'est pas nécessairement valable dans le cas présent, l'extension de la source pouvant ne pas être négligeable devant λ.

[3] Cette définition généralise, en magnétostatique élémentaire, la définition du moment magnétique pour un circuit filiforme $\boldsymbol{\mathcal{M}} = I\mathbf{S}$. On remarquera en effet que l'intégrale de contour de $\mathbf{r}\times d\mathbf{r}/2$ n'est autre que le vecteur surface défini par ce contour.

magnétique oscillant. Si une charge oscillant le long de l'axe Oz modélise un dipôle électrique, une répartition dipolaire magnétique correspondrait plutôt à une petite boucle de courant oscillant (perpendiculaire à l'axe portant $\boldsymbol{\mathcal{M}}_0$). Cette identification donne un sens physique à notre séparation a priori artificielle du potentiel vecteur à l'ordre 1.

On peut écrire avec ces définitions:

$$\mathbf{A}_1^{dm} = \frac{\mu_0}{4\pi}\frac{e^{i(kr-\omega t)}}{r} ik\mathbf{u}_r \times \boldsymbol{\mathcal{M}}_0 \ . \tag{2.20}$$

Pour simplifier les écritures, et sans restreindre la généralité du raisonnement, nous pouvons supposer $\boldsymbol{\mathcal{M}}_0$ aligné avec Oz. Nous aurons alors simplement:

$$\mathbf{A}_1^{dm} = \frac{\mu_0}{4\pi}\frac{e^{i(kr-\omega t)}}{r} ik\mathcal{M}_0(\mathbf{u}_r \times \mathbf{u}_z) \ . \tag{2.21}$$

Nous pouvons établir directement les champs rayonnés à partir de cette expression et de la structure locale d'onde plane. Le champ magnétique s'écrit en effet:

$$\begin{aligned} \mathbf{B}^{dm} &= \boldsymbol{\nabla} \times \mathbf{A} \\ &= ik\mathbf{u}_r \times \mathbf{A} \\ &= \frac{\mu_0}{4\pi}\mathcal{M}_0\frac{e^{i(kr-\omega t)}}{r}k^2(-\sin\theta\mathbf{u}_\theta) \ . \end{aligned} \tag{2.22}$$

La structure locale d'onde plane nous donne alors aussi l'expression du champ électrique du dipôle magnétique:

$$\mathbf{E}^{dm} = \frac{\mu_0}{4\pi}\mathcal{M}_0\frac{e^{i(kr-\omega t)}}{r}ck^2(\sin\theta\mathbf{u}_\phi) \tag{2.23}$$

Cette fois, ce sont les lignes de champ électrique qui sont des cercles d'axe Oz.

On peut aussi procéder par analogie avec le dipôle électrique, ce qui soulignera les similitudes entre les rayonnements de ces deux distributions de charges. Le champ magnétique du dipôle électrique (que nous supposerons également aligné selon Oz) s'écrit en effet:

$$\mathbf{B}^{de} = \frac{\mu_0}{4\pi}\left(\frac{\omega d_0}{i}\right)\frac{e^{i(kr-\omega t)}}{r}ik(\mathbf{u}_r \times \mathbf{u}_z) \ . \tag{2.24}$$

Il est donc, à quelques substitutions de symboles près, identique au potentiel vecteur du dipôle magnétique.

Le champ électrique du dipôle électrique s'obtient en écrivant:

$$\mathbf{E}^{de} = \frac{ic^2}{\omega}\boldsymbol{\nabla} \times \mathbf{B}^{de} \tag{2.25}$$

et le champ magnétique du dipôle magnétique par:

$$\mathbf{B}^{dm} = \boldsymbol{\nabla} \times \mathbf{A}_1^{dm} \ . \tag{2.26}$$

En comparant ces deux expressions, nous voyons que $\mathbf{B}^{dm}$ s'obtient simplement en substituant d_0c^2 par $\mathcal{M}_0$ dans $\mathbf{E}^{de}$. La structure du champ magnétique du dipôle magnétique est donc identique (à ce changement de notation près) à la structure du champ électrique du dipôle électrique. Si ces deux types de sources n'apparaissent pas au même ordre du développement, elles sont tout à fait similaires. En utilisant l'expression (1.78) du champ électrique du dipôle, nous obtenons immédiatement:

$$\mathbf{B}^{dm} = \frac{\mu_0}{4\pi}\mathcal{M}_0\frac{e^{i(kr-\omega t)}}{r}k^2(-\sin\theta\mathbf{u}_\theta) \ . \tag{2.27}$$

Si le terme dipolaire magnétique est le seul à contribuer au rayonnement, nous pouvons estimer le flux d'énergie. Rappelons que si plusieurs termes contribuent, ce calcul ne sera pas complet. Il faudrait aussi tenir compte des termes croisés entre champs à différents ordres du développement multipolaire. Nous ne considérerons donc ici que le cas d'une source dont le dipôle électrique est nul. Nous supposerons aussi que la contribution des termes quadripolaires électriques (la seconde ligne de (2.16)) est négligeable. Le vecteur de Poynting décrivant le flux d'énergie, en valeur moyenne temporelle, s'écrit alors, après quelques transformations élémentaires pour faire apparaître ϵ_0:

$$\overline{\mathbf{\Pi}} = \frac{\mathcal{M}_0^2}{32\pi^2\epsilon_0 c^5}\frac{\omega^4}{r^2}\sin^2\theta\mathbf{u}_r \ . \tag{2.28}$$

Le diagramme de rayonnement du dipôle magnétique est donc identique à celui du dipôle électrique (on se reportera à la figure correspondante). En particulier, le rayonnement est nul dans la direction du moment dipolaire magnétique, maximal dans le plan perpendiculaire (le plan du circuit si on imagine le dipôle sous la forme d'une petite boucle de courant).

La puissance totale s'obtient en intégrant le flux d'énergie sur une sphère de rayon r arbitraire. Le calcul est exactement analogue à celui du dipôle électrique et nous ne le détaillerons pas. La puissance moyenne rayonnée est alors:

$$\mathcal{P}^{dm} = \frac{\mathcal{M}_0^2}{12\pi\epsilon_0 c^5}\omega^4 \ . \tag{2.29}$$

On remarque encore la très rapide dépendance en fréquence de cette puissance. Il est instructif à ce point de tenter de comparer, en ordres de grandeur, la puissance rayonnée par un dipôle électrique et un dipôle magnétique. On peut se poser bien sûr le question du sens de ce problème. Nous avons supposé, pour établir cette expression de la puissance, que le dipôle électrique était strictement nul. On peut cependant imaginer deux répartitions de courant, ayant la même extension spatiale et les mêmes courants caractéristiques, mais des propriétés de symétrie telles que l'une ait un dipôle électrique et l'autre seulement un dipôle magnétique. Si cette démarche peut paraître un peu artificielle pour des répartitions de courant macroscopiques, elle est bien justifiée pour le rayonnement d'atomes. Certaines transitions partagent en effet des fréquences et des "courants" équivalents mais correspondent à des symétries des orbitales mises en jeu telles que l'une se comporte comme un dipôle électrique et l'autre comme un dipôle magnétique. La puissance moyenne dipolaire électrique s'écrit:

$$\mathcal{P}^{de} = \frac{d_0^2}{12\pi\epsilon_0 c^3}\omega^4 \ . \tag{2.30}$$

On en déduit alors

$$\frac{\mathcal{P}^{de}}{\mathcal{P}^{dm}} = \frac{c^2 d_0^2}{\mathcal{M}_0^2} \ . \tag{2.31}$$

Pour aller plus loin, il nous faut "comparer" l'amplitude des dipôles électriques et magnétiques, ce qui n'a de sens qu'en termes d'ordre de grandeur. Nous ne ferons donc qu'un raisonnement très qualitatif. On peut écrire symboliquement:

$$\mathbf{d}_0 = \int \mathbf{r}'\rho_0\, d^3\mathbf{r}' \tag{2.32}$$

$$\boldsymbol{\mathcal{M}}_0 = \int \mathbf{r}'\mathbf{j}_0\, d^3\mathbf{r}' \ . \tag{2.33}$$

Mais $\mathbf{j}_0$ est de l'ordre de $\rho_0\mathbf{v}$ où $\mathbf{v}$ est la vitesse de déplacement des charges au point $\mathbf{r}'$. On voit bien tout le caractère qualitatif de ce raisonnement, qui ne tiendrait que si les deux dipôles étaient décrits précisément par la même distribution de courant, ce qui ne peut être le cas. En ordre de grandeur, néanmoins, on en déduit:

$$\mathcal{M}_0 \approx d_0 v \tag{2.34}$$

où v est un ordre de grandeur de la vitesse maximale des charges. On en déduit donc finalement:

$$\frac{\mathcal{P}^{de}}{\mathcal{P}^{dm}} \approx \left(\frac{c}{v}\right)^2 . \tag{2.35}$$

Notons que ce résultat découle simplement de notre développement en puissances de r_0/λ. Cette quantité est en effet égale à $r_0\omega/c$, de l'ordre de v/c. Le terme dipolaire magnétique est donc, pour ce qui est des champs, v/c fois plus petit que le terme dipolaire électrique, d'où l'ordre de grandeur ci–dessus pour les puissances rayonnées.

Si les charges sont en mouvement à une vitesse faible par rapport à celle de la lumière, le rayonnement dipolaire magnétique est donc beaucoup moins efficace que le rayonnement dipolaire électrique. Si nous considérons par exemple les deux transitions atomiques évoquées plus haut, le rapport des puissances émises est de l'ordre de $(1/\alpha)^2$ où $\alpha = e^2/(4\pi\epsilon_0\hbar c) \approx 1/137$ est la constante de structure fine. Comme nous le verrons dans le chapitre suivant, l'énergie rayonnée doit être empruntée à l'énergie initiale de l'atome et la puissance rayonnée détermine la durée de vie d'un niveau radiatif. Nous voyons donc ainsi que la durée de vie d'un niveau ne pouvant rayonner que comme un dipôle magnétique est environ $1/\alpha^2$ fois plus grande que la durée de vie d'un niveau "dipolaire électrique". Celle-ci étant de quelques nanosecondes (voir chapitre suivant), la durée de vie d'un niveau "dipolaire magnétique", toutes choses égales par ailleurs, est plutôt de la dizaine de microsecondes. Nous verrons à nouveau une illustration de cette différence entre dipôles électriques et magnétiques quand nous considérerons des rayonnements d'antennes. Notons pour finir que ce raisonnement prouve, s'il en était besoin, que le terme dipolaire électrique domine le terme dipolaire magnétique s'ils sont simultanément présents pour une même distribution de courants.

Quadripôle électrique.

Nous allons maintenant traiter le terme $\mathbf{A}_1^{qe}$ correspondant à la seconde ligne de l'équation (2.16). Il s'écrit donc

$$\mathbf{A}_1^{qe} = \frac{\mu_0}{4\pi}\frac{e^{i(kr-\omega t)}}{r}(-ik)\int \frac{1}{2}\left[(\mathbf{u}_r\cdot\mathbf{r}')\mathbf{j}_0 + \mathbf{r}'(\mathbf{u}_r\cdot\mathbf{j}_0)\right]\, d^3\mathbf{r}' \tag{2.36}$$

Nous allons essayer de transformer l'intégrande pour le mettre sous une forme plus agréable. En particulier, nous allons tenter de faire apparaître la densité de charge (nous verrons que le moment quadripolaire électrique est au moment dipolaire ce que le tenseur d'inertie est au centre de gravité). Pour effectuer ce calcul sans trop de peine, nous abandonnerons un temps les notations de l'analyse vectorielle standard. Nous aurons en effet à manipuler des quantités dyadiques (tensorielles de rang deux, en d'autres termes). Nous expliciterons donc les dérivées, en utilisant les conventions d'Einstein pour alléger les calculs.

L'intégrande apparaissant dans le potentiel vecteur s'écrit avec ces conventions:

$$\frac{1}{2}[u_j r'_j j_i + r'_i u_j j_j] . \tag{2.37}$$

L'intégration se faisant uniquement sur $\mathbf{r}'$, on peut donc écrire:

$$\int \frac{1}{2}\left[(\mathbf{u}_r\cdot\mathbf{r}')\mathbf{j}_0 + \mathbf{r}'(\mathbf{u}_r\cdot\mathbf{j}_0)\right]\, d^3\mathbf{r}'\Big]_i = \frac{1}{2}(-i\omega)\mathcal{Q}_{ij}u_j , \tag{2.38}$$

avec

$$-i\omega\mathcal{Q}_{ij} = \int [j_i r'_j + j_j r'_i]\, d^3\mathbf{r}' . \tag{2.39}$$

Le facteur $-i\omega$ a été introduit ici pour simplifier les calculs suivants. Pour expliciter le tenseur $\mathcal{Q}$, remarquons que:

$$\partial'_k(r'_i r'_j j_k) = \delta_{ki} r'_j j_k + \delta_{kj} r'_i j_k + r'_i r'_j (i\omega\rho) , \tag{2.40}$$

où toutes les dérivées spatiales s'entendent par rapport à $\mathbf{r}'$. Nous avons utilisé l'équation de conservation de la charge, qui s'écrit ici $\partial'_k j_k = i\omega\rho_0$. On en tire immédiatement:

$$\int (j_i r'_j + j_j r'_i)\, d^3\mathbf{r}' = -i\omega \int (r'_i r'_j \rho_0)\, d^3\mathbf{r}' \tag{2.41}$$

puisque l'intégrale de $\partial'_k(r'_i r'_j j_k)$, intégrale d'une divergence sur tout l'espace, est manifestement nulle. Finalement, on a:

$$\mathcal{Q}_{ij} = \int r'_i r'_j \rho_0(\mathbf{r}')\, d^3\mathbf{r}' \ , \tag{2.42}$$

que nous appellerons tenseur quadripolaire électrique (ce tenseur est évidemment symétrique). Nous noterons en effet que cette définition coïncide avec celle du quadripôle dans le cadre de l'électrostatique et soulignerons l'analogie avec le tenseur d'inertie en mécanique du solide.

Avec cette définition, les composantes du potentiel vecteur s'écrivent:

$$A^{qe}_{1i} = -\frac{\mu_0}{4\pi}\frac{\omega^2}{2c}\frac{e^{i(kr-\omega t)}}{r}\mathcal{Q}_{ij}u_{rj} \ . \tag{2.43}$$

Plutôt que de conserver des expression faisant intervenir les composantes, nous pouvons définir le vecteur $\boldsymbol{\mathcal{Q}}\cdot\mathbf{u}_r$ comme le vecteur de composantes $u_{rj}\mathcal{Q}_{ji}$. Le potentiel vecteur s'écrivant alors:

$$\mathbf{A}^{qe}_1 = -\frac{\mu_0}{4\pi}\frac{\omega^2}{2c}\frac{e^{i(kr-\omega t)}}{r}\boldsymbol{\mathcal{Q}}\cdot\mathbf{u}_r \ . \tag{2.44}$$

Avant de discuter les propriétés des champs rayonnés, quelques remarques s'imposent. D'abord, nous avons ici introduit le quadripôle électrique de la manière la plus simple mais pas de la manière la plus commode pour les développements mathématiques. Il est fructueux de définir le quadripôle comme un tenseur symétrique de trace nulle (on consultera à ce sujet le Jackson). La définition "standard" du tenseur est en fait:

$$\mathcal{Q}'_{ij} = 3\mathcal{Q}_{ij} - \delta_{ij}\int r'^2\rho_0\, d^3\mathbf{r}' \tag{2.45}$$

dont la trace est évidemment nulle. A part un facteur trois, qui n'introduit que des différences algébriques mineures dans les expressions, ces deux définitions diffèrent par un terme diagonal dont les trois éléments sont égaux. Le potentiel vecteur correspondant à ce terme diagonal est manifestement colinéaire à $\mathbf{u}_r$. Son rotationnel est nul et les deux définitions conduisent aux mêmes champs rayonnés. Nous n'adopterons pas la convention la plus générale. En effet, nous ne traiterons guère en détails le rayonnement du quadripôle : il ne sera donc pas très gênant de travailler avec un tenseur de trace non nulle. Nous aurons en revanche intérêt à travailler avec la forme la plus simple possible.

Essayons d'imaginer maintenant des distributions de charge qui ne possèdent pas de dipôles électriques mais un quadripôle électrique. La plus simple, comme en électrostatique, est constituée de deux dipôles électriques oscillant en opposition de phase (amplitudes opposées) et placés à petite distance l'un de l'autre. Les deux contributions dipolaires s'annulent et il ne reste qu'une contribution quadripolaire que l'on pourra calculer à titre d'exercice. Une autre distribution quadripolaire est constitué d'une répartition de charge uniforme dont la forme évolue entre un ellipsoïde allongé selon Oz et un ellipsoïde aplati dans la même direction, en passant par la forme sphérique à mi–période. Pour des raisons de symétrie évidentes (symétrie par rapport au plan xOy), cette distribution ne possède de moment dipolaire ni électrique ni magnétique. En raisonnant par analogie avec le moment d'inertie, on voit bien que cette distribution possède en revanche une composante oscillante non nulle du quadripôle électrique. Notons enfin, comme nous le verrons dans un prochain paragraphe, qu'une charge ponctuelle oscillant le long de Oz possède, en plus de son moment dipolaire électrique, un moment quadripolaire si l'extension du mouvement n'est pas négligeable.

Nous pouvons maintenant calculer les champs rayonnés. Nous utiliserons pour cela la structure locale d'onde plane. On trouve ainsi immédiatement:

$$\mathbf{B}^{qe} = -\frac{\mu_0}{4\pi} i \frac{\omega^3}{2c^2} \frac{e^{i(kr-\omega t)}}{r} \mathbf{u}_r \times (\boldsymbol{\mathcal{Q}} \cdot \mathbf{u}_r) \tag{2.46}$$

Le champ électrique s'en déduisant sans difficultés.

Avec les mêmes restrictions que pour le dipôle magnétique, nous pouvons, dans le cas ou le rayonnement est purement quadripolaire électrique, écrire enfin le vecteur de Poynting en valeur moyenne temporelle:

$$\overline{\mathbf{\Pi}}^{qe} = \frac{1}{128\pi^2\epsilon_0 c^5} \omega^6 \frac{|\mathbf{u}_r \times (\boldsymbol{\mathcal{Q}} \cdot \mathbf{u}_r)|^2}{r^2} \tag{2.47}$$

Nous noterons que la dépendance en ω est encore plus rapide que pour les rayonnements dipolaires.

Il est bien difficile de préciser davantage la structure de ces champs sans faire quelques hypothèses sur la structure du quadripôle électrique. Nous supposerons donc que le seul élément non nul est $\mathcal{Q}_{zz} = Q_0$. C'est le cas de la composante quadripolaire pour la charge oscillante. Pour la répartition ellipsoïdale, la symétrie impose seulement que les termes non diagonaux soient nuls. Elle impose aussi $Q_{xx} = Q_{yy}$. On peut alors retirer au quadripôle un terme diagonal ayant trois valeurs propres égales à Q_{xx}. Nous savons que cela ne change pas les champs rayonnés. En revanche, nous pouvons ainsi annuler tous les termes de la diagonale sauf Q_{zz}. La situation que nous décrivons maintenant, sans être tout à fait générale, n'en est pas moins importante.

Dans ce cas, nous avons simplement $\mathbf{u}_r \cdot \boldsymbol{\mathcal{Q}} = Q_0 \cos\theta \mathbf{u}_z$. Nous pouvons alors réécrire sans difficultés (en utilisant $\mathbf{u}_r \times \mathbf{u}_z = -\sin\theta \mathbf{u}_\phi$) potentiel vecteur, champ magnétique et vecteur de Poynting moyens comme:

$$\mathbf{A}_1^{qe} = -\frac{\mu_0}{4\pi} \frac{\omega^2}{2c} \frac{e^{i(kr-\omega t)}}{r} Q_0 \cos\theta \mathbf{u}_z \tag{2.48}$$

$$\mathbf{B}_1^{qe} = i\frac{\mu_0}{4\pi} \frac{\omega^3}{2c^2} \frac{e^{i(kr-\omega t)}}{r} Q_0 \cos\theta \sin\theta \mathbf{u}_\phi \tag{2.49}$$

$$\overline{\mathbf{\Pi}}^{qe} = \frac{1}{128\pi^2\epsilon_0 c^5} \omega^6 Q_0^2 \frac{\cos^2\theta \sin^2\theta}{r^2} \mathbf{u}_r \tag{2.50}$$

Comme pour le dipôle électrique, le champ magnétique est selon $\mathbf{u}_\phi$ et ses lignes sont des cercles d'axe Oz. Le champ électrique est dans le plan défini par l'axe du quadripôle et la direction d'observation. Le diagramme de rayonnement du quadripôle est substantiellement différent de celui du dipôle. La puissance rayonnée s'annule aussi bien dans la direction Oz que sur le plan xOy. Le diagramme de rayonnement, en coupe dans un plan passant par Oz, se présente donc comme quatre lobes symétriques autour de l'origine[4]. Ce diagramme de rayonnement est représenté sur la figure 2.2.

Finalement, on peut obtenir la puissance totale rayonnée en intégrant le vecteur de Poynting sur une sphère de rayon r arbitraire (la dépendance en r disparaissant évidemment). L'intégration sur ϕ est triviale et donne un simple facteur 2π. On trouve donc:

$$\mathcal{P}^{qe} = \frac{1}{64\pi\epsilon_0 c^5} \omega^6 Q_0^2 \int_0^\pi \sin^3\theta \cos^2\theta \, d\theta \ . \tag{2.51}$$

L'intégrale sur θ est facilement évaluée et vaut 4/15. On a donc finalement:

$$\mathcal{P}^{qe} = \frac{\omega^6 Q_0^2}{240\pi\epsilon_0 c^5} \tag{2.52}$$

Nous noterons encore une fois la dépendance en ω^6 de la puissance émise. On peut comprendre qualitativement cette dépendance de plus en plus rapide. Plus nous progressons dans les ordres

[4]En termes d'harmoniques sphériques, le dipôle est associé à celle de moment cinétique total 1 Y^1, qui présente deux lobes symétriques par rapport à l'origine, le quadripôle à l'harmonique sphérique de moment 2 Y^2

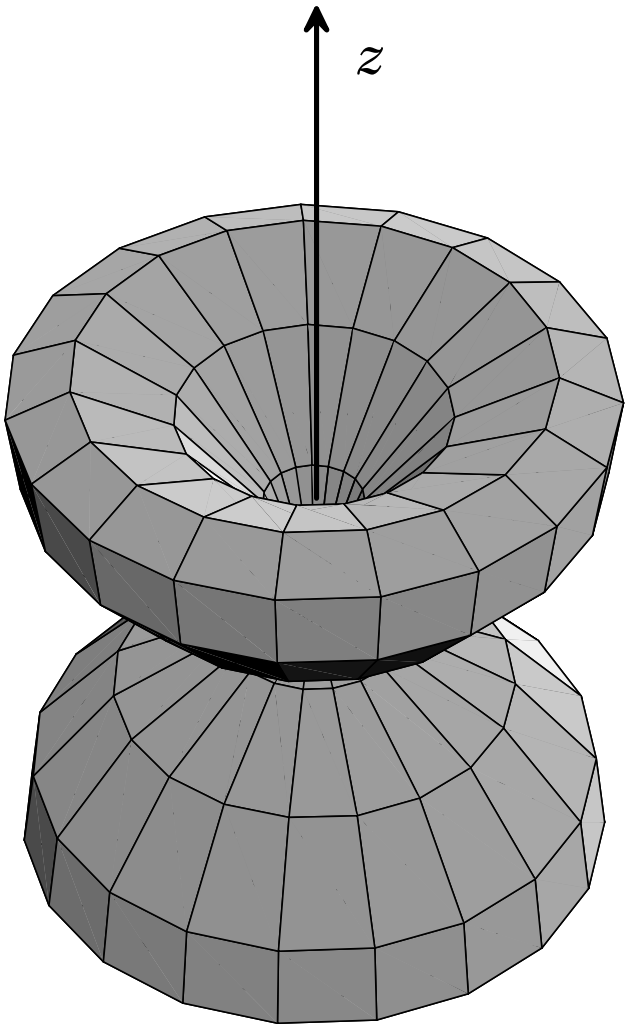

Figure 2.2: Diagramme du rayonnement d'un quadripôle n'ayant comme seul élément non nul que Q_{zz}. On note que le rayonnement est nul sur l'axe et dans un plan perpendiculaire à l'axe.

multipolaires, plus nous considérons des modulations subtiles de la répartition de charges. Toutes choses égales par ailleurs, plus la longueur d'onde rayonnée sera courte, plus les champs seront sensibles à ces très faibles dissymétries de la répartition de charges.

Comme nous l'avons fait pour le dipôle magnétique, et avec les mêmes restrictions sur la validité du calcul, nous pouvons maintenant comparer, toutes choses égales par ailleurs, les puissances dipolaires et quadripolaires électriques. On trouve immédiatement:

$$\frac{\mathcal{P}^{qe}}{\mathcal{P}^{de}} = \frac{\omega^2 Q_0^2}{d_0^2} \frac{1}{20c^2} \ . \tag{2.53}$$

Dans un raisonnement en termes d'ordres de grandeur, on peut approcher d_0 par le produit qr_0 de la charge totale par l'extension spatiale des courants. Dans ce cas, Q_0 est de l'ordre de qr_0^2. On a donc finalement:

$$\frac{\mathcal{P}^{qe}}{\mathcal{P}^{de}} = \frac{k^2 r_0^2}{20} \ . \tag{2.54}$$

A part le facteur numérique 1/20, cette expression était tout à fait prévisible. N'oublions pas que les termes dipolaires électriques et quadripolaires électriques apparaissent comme deux termes successifs dans un développement en puissances de l'extension de la source rapportée à λ. Les champs sont, à chaque ordre, multipliés par un facteur kr_0 et les puissances par un facteur $k^2 r_0^2$. Le rayonnement quadripolaire est donc beaucoup moins efficace que le rayonnement dipolaire si l'extension de la source est très petite devant λ. Si nous considérons encore deux niveaux atomiques se désexcitant par des transitions en tous points comparables, l'une dipolaire et l'autre quadripolaire, le rapport des puissances émises et donc des durées de vie sera de l'ordre de $(a_0/\lambda)^2$, où a_0 est le rayon de Bohr. Ce rapport est numériquement de l'ordre de 10^{-8}. Les niveaux quadripolaires électriques ont une durée de vie 10^7 à 10^9 fois plus grande que ceux qui se désexcitent comme un dipôle électrique. Les durées de vie de ces niveaux "métastables" peuvent atteindre la fraction de seconde. La raie associée à ces transitions est alors très fine, fort intéressante pour la métrologie de fréquence dans le domaine optique.

Nous pourrions poursuivre encore plus loin le développement multipolaire. A l'ordre suivant, nous pourrions distinguer deux contributions: le quadripôle magnétique (modélisé par exemple par deux spires parcourues par des courants oscillants en opposition de phase) dont le rayonnement est très similaire à celui du quadripôle électrique et l'octupôle électrique (deux quadripôles voisins en opposition de phase). En fait, l'étude systématique des moments dipolaires d'ordre élevé doit être entreprise avec des techniques dérivées de la théorie des groupes.

2.3 Applications: quelques problèmes de rayonnement

Nous allons dans ce paragraphe appliquer les développements multipolaires à l'étude du rayonnement de quelques sources simples. Nous commencerons par étudier le rayonnement d'une charge oscillante en ne faisant plus l'hypothèse, comme au chapitre précédent, que l'amplitude du mouvement est négligeable. Nous pourrons ainsi préciser la validité des hypothèses utilisées pour le rayonnement du dipôle et voir quelles corrections il y a lieu d'apporter quand ces hypothèses ne sont plus valables. Nous aborderons très brièvement ensuite le très riche problème des antennes. Il est si vaste — des manuels entiers lui sont consacrés — que nous ne pourrons traiter très brièvement que deux exemples: l'antenne dipolaire électrique et l'antenne dipolaire magnétique.

2.3.1 Rayonnement d'une charge oscillante

Les notations et le problème sont ceux du dernier paragraphe du chapitre précédent. Une charge q est animée d'un mouvement harmonique d'amplitude a et de fréquence ω le long de l'axe Oz. Nous ne nous préoccupons que du champ rayonné. Tous les résultats précédents étant acquis, le seul problème

est de déterminer les différents éléments de la source: dipôle électrique et magnétique, quadripôle électrique.

De façon rigoureuse, le courant associé à la particule peut s'écrire:

$$\mathbf{j}(\mathbf{r}') = -qa\omega \sin\omega t \delta(x')\delta(y')\delta(z' - a\cos\omega t)\mathbf{u}_z \ . \tag{2.55}$$

C'est simplement le produit de la charge par la vitesse de la particule, localisé à la position de la particule. Notons que nous utilisons ici, pour un temps, des notations réelles pour les quantités oscillantes et que nous nous intéressons au courant complet, pas seulement à une amplitude d'oscillation. Dans le chapitre précédent, nous avions supposé le courant localisé à l'origine et remplacé donc la dernière fonction δ par $\delta(z')$. Pour aller plus loin, nous pouvons réaliser un développement de Taylor de cette dernière fonction δ dont les deux premiers termes nous donneront l'essentiel de la physique si l'étendue du mouvement n'est pas trop grande. Nous écrirons donc:

$$\delta(z' - a\cos\omega t) = \delta(z') - a\cos\omega t \delta'(z') + \dots \ . \tag{2.56}$$

Un tel développement de Taylor d'une "fonction" aussi singulière peut paraître hasardeux. On peut se rassurer un peu en montrant l'identité de l'action de la distribution initiale et du développement sur une fonction régulière $f(z')$. Avec les notations de la théorie des distributions, nous avons en effet:

$$\langle \delta(z' - a\cos\omega t)|f\rangle = f(a\cos\omega t) = f(0) + f'(0)a\cos\omega t + \dots \ , \tag{2.57}$$

puisque f, régulière en zéro, peut être développée en série de Taylor. Sachant que $\langle \delta'|f\rangle = -f'(0)$, on reconnaît sans difficulté dans le second membre l'action de la distribution "développée". En nous limitant aux deux premiers termes, nous écrirons donc

$$\mathbf{j} = -qa\omega \sin\omega t \delta(\mathbf{r}') + qa^2\omega \sin\omega t \cos\omega t \delta(x')\delta(y')\delta'(z') \ . \tag{2.58}$$

Dans le premier terme du second membre, nous reconnaissons bien sûr le courant localisé à l'origine que nous avions utilisé en première approximation. Le second terme, lui, est clairement un courant oscillant à la fréquence 2ω. Contrairement à ce qu'on aurait pu penser de prime abord, les termes correctifs ne sont pas à la fréquence fondamentale, mais à ses harmoniques.

Nous écrirons donc finalement:

$$\mathbf{j} = \mathbf{j}_1 e^{-i\omega t} + \mathbf{j}_2 e^{-2i\omega t} \ , \tag{2.59}$$

avec

$$\mathbf{j}_1 = -iqa\omega\delta(\mathbf{r}')\mathbf{u}_z \tag{2.60}$$

$$\mathbf{j}_2 = i\frac{qa^2\omega}{2}\delta(x')\delta(y')\delta'(z')\mathbf{u}_z \tag{2.61}$$

et nous examinerons brièvement ces deux termes.

$\mathbf{j}_1$ est le courant du dipôle. Le moment dipolaire magnétique ferait intervenir des intégrales de $x\delta(x)$ (ou de cette forme) qui sont identiquement nulles. Pour estimer le dipôle et le quadripôle électriques, il nous faut la densité de charges. A partir de l'équation de continuité, nous trouvons:

$$\rho_1(\mathbf{r}') = -qa\delta(x')\delta(y')\delta'(z') \tag{2.62}$$

Les composantes x et y du dipôle électrique font intervenir aussi des intégrales de type $x\delta(x)$ qui sont nulles. Seule la composante selon z est non nulle et le dipôle s'écrit (en notant que $\int z\delta'(z) = -1$):

$$\mathbf{d}_1 = qa\mathbf{u}_z \ , \tag{2.63}$$

un résultat que nous aurions certainement pu prédire directement. Toutes les composantes du tenseur quadripôle électrique qui font intervenir une composante x ou y sont nulles. Elles contiennent en

effet des intégrales de $x\delta(x)$ ou $x^2\delta(x)$ qui sont nulles. La composante Q_{zz} est proportionnelle à $\int z'^2\delta'(z')\,dz'$. Elle est également nulle. Ce courant est donc un pur dipôle électrique (le même genre d'arguments peut être utilisé pour montrer que tous les autres termes du développement multipolaire sont nuls).

$\mathbf{j}_2$ ne peut, lui non plus, correspondre à un dipôle magnétique. Les moments dipolaire et quadripolaire s'obtiennent à partir de la densité de charges qui s'écrit:

$$\rho_2 = -\frac{i}{2\omega}\boldsymbol{\nabla}\cdot\mathbf{j}_2 = \frac{qa^2}{4}\delta(x')\delta(y')\delta''(z') \ . \tag{2.64}$$

Comme pour ρ_1, les composantes x et y du dipôle sont manifestement nulles. La composante z fait intervenir l'intégrale $\int z'\delta''(z')$ qui est nulle, comme on pourra s'en convaincre par une intégration par parties. Pour le quadripôle, toutes les composantes faisant intervenir x ou y sont identiquement nulles. Seule la composante Q_{2zz} peut être non nulle. On a effectivement:

$$Q_{2zz} = \frac{qa^2}{4}\int z'^2\delta''(z')\,dz' \ . \tag{2.65}$$

Une double intégration par parties prouve que $\int z'^2\delta''(z')\,dz' = 2$. On a donc finalement:

$$Q_{2zz} = \frac{qa^2}{2} \ . \tag{2.66}$$

Le courant $\mathbf{j}_2$ correspond donc à un pur quadripôle électrique (on pourrait montrer qualitativement que tous les termes d'ordre supérieur sont identiquement nuls).

Les champs dipolaires électriques et quadripolaires électriques étant à des fréquences différentes, il n'y a pas, dans le calcul de la puissance totale rayonnée en valeur moyenne temporelle, de termes croisés et on peut écrire $\mathcal{P} = \mathcal{P}^{de} + \mathcal{P}^{qe}$. On peut remarquer que, conformément aux ordres de grandeur qualitatifs des paragraphes précédents,

$$\frac{\mathcal{P}^{qe}}{\mathcal{P}^{de}} = \frac{k^2a^2}{80} \tag{2.67}$$

On vérifie ainsi que notre première approximation, consistant à négliger la composante quadripolaire électrique à 2ω est largement justifiée, même pour des extensions du mouvement voisines de λ.

2.3.2 Antennes

Le problème des antennes est d'un importance technologique majeure. Il s'agit en effet, pour optimiser les communications, de rayonner la plus grande puissance possible, avec le meilleur rendement vis à vis de la puissance fournie par les générateurs. On cherche souvent aussi à optimiser les propriétés de polarisation de la lumière émise. Pour une propagation au voisinage de la surface terrestre, conductrice, une polarisation verticale est indispensable. On peut enfin chercher à optimiser le diagramme de rayonnement, avec des réseaux d'antennes par exemple, pour que toute la puissance émise le soit dans la direction utile. Le problème est si vaste qu'il est hors de question de le traiter ici de manière complète. De plus, la caractérisation complète d'une antenne de géométrie donnée est un problème formidablement complexe. Déterminer le rayonnement, c'est d'abord déterminer les courants dans l'antenne. Ces courants dépendent des générateurs mais aussi des champs produits par l'antenne elle même. Le courant doit donc a priori être déterminé de façon "self consistante" en tenant compte de cette réaction de rayonnement. On peut utiliser pour cela une résolution explicite des équations de Maxwell (éventuellement numérique) ou des méthodes de perturbations. Nous nous contenterons ici de la première étape de ces calculs en postulant une forme simple pour les courants dans l'antenne.

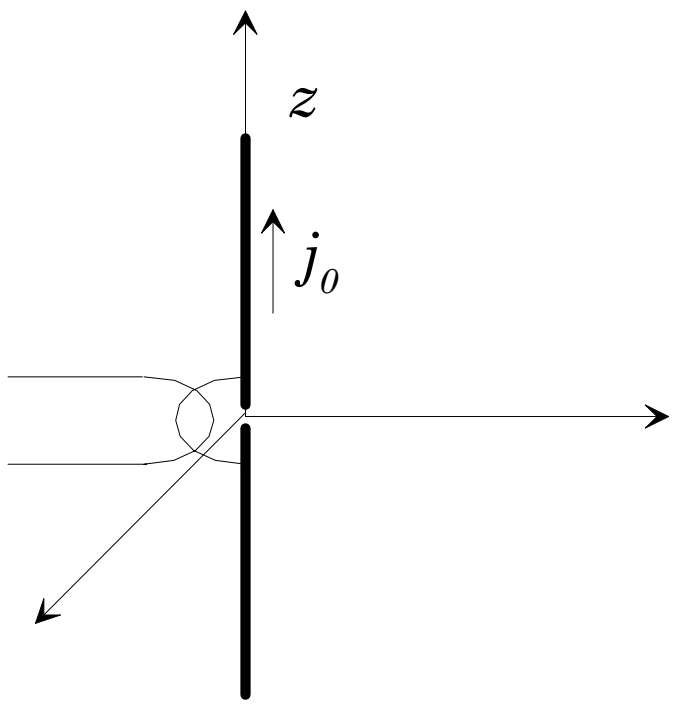

Figure 2.3: Antenne dipolaire électrique

Antenne dipolaire électrique

La première antenne que nous considérerons est constituée d'un simple fil rectiligne, de longueur ℓ, aligné avec l'axe Oz, symétrique par rapport à O. Ce genre d'antenne est effectivement très largement utilisé, tant en réception qu'en émission. Dans la pratique, ce genre d'antennes est attaqué par un secondaire de transformateur haute fréquence situé en série avec l'antenne au voisinage de O. Le courant circule donc alternativement dans les directions positives et négatives de l'axe. De manière évidente, le courant s'annule aux extrémités de l'antenne. Il est assez naturel et sans doute assez réaliste de postuler une répartition de courant sinusoïdale le long de l'antenne et d'écrire l'amplitude complexe du courant:

$$\mathbf{j}_0(\mathbf{r}') = I_0 \cos \frac{\pi z'}{\ell} \delta(x')\delta(y')\mathbf{u}_z \ , \tag{2.68}$$

pour $-\ell/2 \leq z' \leq \ell/2$, le courant étant nul ailleurs. Les deux fonctions δ servant à localiser le courant sur l'axe. Il est évident dès l'abord qu'une telle distribution de courant ne possède pas de moment dipolaire magnétique (ce que l'on pourra vérifier par un calcul explicite). Si le moment dipolaire est non nul, c'est lui qui dominera le rayonnement. Pour l'estimer, nous allons calculer la densité de charges:

$$\rho_0 = -\frac{i}{\omega} \boldsymbol{\nabla} \cdot \mathbf{j}_0 = \frac{iI_0}{\omega} \frac{\pi}{\ell} \sin \frac{\pi z'}{\ell} \delta(x')\delta(y') \ . \tag{2.69}$$

Les composantes selon x et y du moment dipolaire sont nulles, puisqu'elles font intervenir des intégrales du type $x'\delta(x')$. Le dipôle est aligné selon Oz, comme l'indique la symétrie de l'antenne. Il vaut:

$$\mathbf{d}_0 = d_0 \mathbf{u}_z \ , \tag{2.70}$$

avec

$$\begin{aligned} d_0 &= \frac{iI_0}{\omega} \frac{\pi}{\ell} \int_{-\ell/2}^{\ell/2} z' \sin \frac{\pi z'}{\ell} \, dz' \\ &= \frac{2iI_0}{\omega} \frac{\ell}{\pi} \end{aligned} \tag{2.71}$$

A moins que l'antenne ne soit très longue par rapport à λ, le rayonnement dipolaire domine et la puissance moyenne rayonnée vaut:

$$\mathcal{P} = \frac{1}{12\pi\epsilon_0 c^3}\omega^4 \left(\frac{2I_0\ell}{\pi\omega}\right)^2 = \frac{I_0^2\omega^2\ell^2}{3\pi^3\epsilon_0 c^3} \ . \tag{2.72}$$

Cette puissance est directement proportionnelle au carré de l'amplitude du courant qui passe dans l'antenne. Il est donc naturel de comparer cette puissance rayonnée à la puissance qui serait dissipée, en valeur moyenne, par le même courant circulant dans une résistance. La résistance équivalente à l'antenne, encore appelée résistance de rayonnement, $R_{\rm rad}$, est finalement définie par:

$$\mathcal{P} = \frac{I_0^2}{2} R_{\rm rad} \ . \tag{2.73}$$

Après quelques manipulations, on trouve:

$$R_{\rm rad} = \frac{8}{3\pi\epsilon_0 c}\left(\frac{\ell}{\lambda}\right)^2 \tag{2.74}$$

$$= 319\left(\frac{\ell}{\lambda}\right)^2 \ \Omega \ . \tag{2.75}$$

La résistance de rayonnement est une quantité très importante pour caractériser une antenne. Si on désire assurer la meilleure utilisation de la puissance fournie par les générateurs, il faut en effet réaliser une *adaptation d'impédance* entre le générateur et l'antenne: le rapport de la puissance dissipée dans l'antenne à la puissance dissipée dans la résistance interne du générateur est maximum quand la résistance équivalente de l'antenne est égale à la résistance interne, ou plutôt à l'impédance de sortie, du générateur. Pour une antenne dite "demi-onde", par exemple, telle que $\ell = \lambda/2$, la résistance de rayonnement vaut $319/4 = 79\ \Omega$. On comprend ainsi pourquoi les impédances de sortie des amplificateurs et les impédances itératives des câbles coaxiaux utilisés pour le branchement des antennes sont de 75 Ω. Notons que l'antenne demi-onde est très favorable, la longueur du fil étant alors parfaitement adaptée à une structure d'onde stationnaire du courant sur l'antenne. En fait, l'antenne toute entière est une structure résonnante, au voisinage d'une fréquence correspondant à $\ell = \lambda/2$, avec des nœuds de courant aux deux extrémités et un ventre au centre. Une antenne quart d'onde, elle, serait résonnante à condition d'être attaquée par le transformateur à une extrémité (on aurait alors un ventre de courant à cette extrémité et un nœud à l'autre). On mesure là l'efficacité des antennes dipolaires électriques: un courant de $0.15A$ suffit à rayonner une puissance de l'ordre du Watt. Le calcul que nous avons fait ici est bien sûr fondé sur un modèle particulier de la répartition de courant dans l'antenne qui pourrait s'avérer inexact. On pourra, à titre d'exercice, faire le calcul pour d'autres répartitions, par exemple pour un courant variant linéairement le long de l'antenne. On s'apercevra alors que seul change le facteur numérique devant la résistance de rayonnement, l'ordre de grandeur restant le même.

Antenne dipolaire magnétique

La deuxième antenne que nous considérerons est représentée sur la figure 2.4. Il s'agit d'une simple spire de courant circulaire, de rayon r_0, située dans le plan xOy. Le courant, amené par deux fils parallèles (que l'on pourra supposer confondus et qui ne jouent donc aucun rôle dans le rayonnement), sera supposé constant sur toute la spire et égal à $I_0 \exp(-i\omega t)$ (c'est une hypothèse vraisemblable si la longueur de la spire est petite devant la longueur d'onde). En coordonnées sphériques, la densité de courant pourra s'écrire:

$$\mathbf{j}_0 = (I_0/r_0)\delta(r' - r_0)\delta(\theta - \pi/2)\mathbf{u}_\phi \ . \tag{2.76}$$

Les fonctions δ localisent bien le courant sur le cercle. Le facteur r_0 est introduit pour donner à j_0 la dimension d'un courant par unité de surface (rappelons que la fonction $\delta(r' - r_0)$ à la dimension

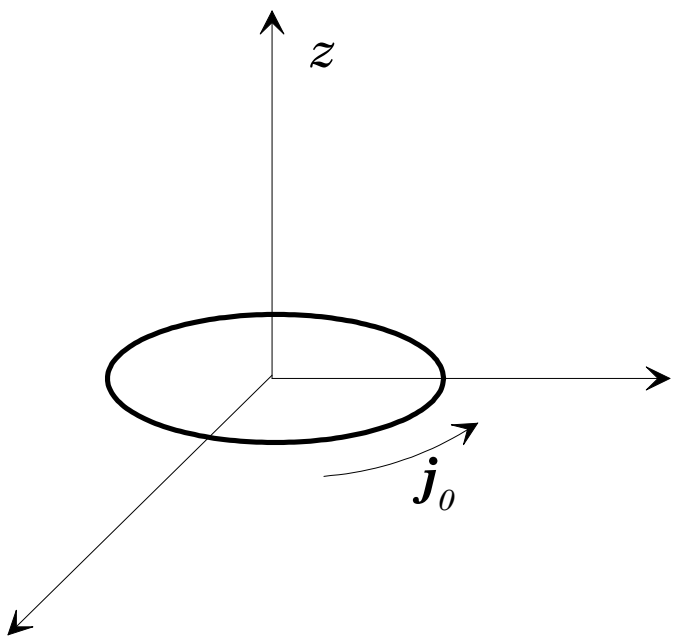

Figure 2.4: Antenne dipolaire magnétique

de l'inverse d'une longueur, alors que la fonction de dirac de l'angle est sans dimension). Finalement, comme on pourra s'en convaincre aisément, ce facteur r_0 assure bien que l'intensité totale circulant dans le circuit ait une amplitude I_0.

Pour estimer les moments dipolaires et quadripolaires électriques, il nous faut la densité de charges. Il est évident que la divergence du courant est nulle: il est orienté selon $\mathbf{u}_\phi$ et ne dépend pas de ϕ. Il n'y a donc aucun moment multipolaire de nature électrique à aucun ordre. S'il n'est pas nul, le moment dipolaire magnétique dominera le rayonnement. On peut l'écrire facilement:

$$\boldsymbol{\mathcal{M}}_0 = \frac{1}{2}\int \mathbf{r}' \times \mathbf{j}_0 \, d^3\mathbf{r}' \ . \tag{2.77}$$

En remarquant que $\mathbf{r}' \times \mathbf{j}_0$ est toujours orienté selon $\mathbf{u}_z$, on voit qu'il en est de même pour le moment magnétique. On peut donc écrire la composante z du moment sous la forme:

$$\mathcal{M}_0 = \frac{1}{2}\int r^3 dr \sin\theta d\theta d\phi \frac{I_0}{r_0}\delta(r' - r_0)\delta(\theta - \pi/2) \ . \tag{2.78}$$

L'intégrale sur ϕ donne un facteur 2π, l'intégrale sur r_0 revient simplement à faire $r' = r_0$. Finalement, l'intégrale sur θ donne 1. On a donc, sans surprises:

$$\boldsymbol{\mathcal{M}}_0 = \pi r_0^2 I_0 \mathbf{u}_z \ . \tag{2.79}$$

Il est facile à partir de là de calculer la puissance rayonnée et la résistance de rayonnement, définie comme au paragraphe précédent. On trouve:

$$\mathcal{P} = \frac{\pi r_0^4 I_0^2 \omega^4}{12\epsilon_0 c^5} \tag{2.80}$$

$$R_{\text{rad}} = \frac{8\pi^5}{2\epsilon_0 c}\left(\frac{r_0}{\lambda}\right)^4 \tag{2.81}$$

$$= 160\left(\frac{r_0}{\lambda}\right)^4 \ \text{k}\Omega \ . \tag{2.82}$$

La résistance de rayonnement croît beaucoup plus vite avec la taille que pour l'antenne dipolaire électrique. A taille égale, l'antenne magnétique est beaucoup plus résistive. Elle est donc plus difficile à exciter, ce qui explique son faible emploi sauf pour des communications à très basses fréquences. On peut comprendre qualitativement cette forte résistance. Si l'antenne dipolaire était équivalente à une très faible inductance (celle du fil) en série avec une très faible capacité (encore celle du fil), l'antenne magnétique présente une inductance élevée, la rendant très résistive à haute fréquence.

Chapitre 3

Sources atomiques de rayonnement

Nous allons dans ce chapitre appliquer les résultats précédents au rayonnement des sources fondamentales que sont les atomes. En fait, nous nous poserons essentiellement deux problèmes. Le premier est celui de l'émission spontanée. Un atome excité, par absorption d'un photon ou collision électronique dans une décharge, passe dans un niveau d'énergie excité. Un simple examen d'une lampe au sodium nous apprend qu'il réémet cette énergie sous forme lumineuse. Est-il possible de préciser les constantes de temps de ce phénomène, c'est à dire la durée de vie des niveaux atomiques excités? L'autre problème, d'une importance pratique aussi considérable, est celui de la diffusion de lumière par les atomes. Quand un atome est placé dans une onde incidente, il acquiert un dipôle à la même fréquence et rayonne. C'est ce phénomène de diffusion qui est, comme nous l'avons déjà mentionné plus haut, à l'origine du bleu du ciel. Peut on comprendre cette diffusion atomique?

Nous essaierons d'abord de comprendre émission spontanée et diffusion dans un modèle complètement classique de l'atome, le modèle de Thomson de l'électron élastiquement lié. Chaque atome est alors assimilé à un oscillateur harmonique chargé ne possédant qu'une fréquence de résonance. C'est bien sûr un modèle très naïf comparé à ce que nous savons de la structure atomique. En revanche, sa simplicité nous permettra de calculer explicitement une durée de vie radiative et de traiter complètement le problème de la diffusion. Nous verrons en particulier comment l'efficacité de diffusion varie quand l'onde excitatrice passe à résonance avec la fréquence propre de l'oscillateur. De manière surprenante, nous verrons que les ordres de grandeur déduits de ce modèle paraissent tout à fait convenables.

Dans le deuxième paragraphe, nous tenterons de donner un modèle semi–classique. Nous traiterons l'atome de manière quantique (nous nous limiterons, pour des raisons évidentes de simplicité, à des structures hydrogénoïdes). Nous comprendrons rapidement que l'émission spontanée ne peut être décrite dans ce cadre. En l'absence de tout rayonnement incident, les niveaux stationnaires ont bien sûr une durée de vie infinie. En fait, l'émission spontanée ne peut être traitée rigoureusement que dans un modèle complètement quantique, où le champ, comme l'atome, est quantifié. Ce modèle inclut en particulier les fluctuations de point zéro du champ électromagnétique, souvent appelées "fluctuations du vide", et prédit l'instabilité des niveaux atomiques autres que le fondamental (qui, lui, fort heureusement, reste stable). Nous nous contenterons donc de traiter le problème de la diffusion de rayonnement. Nous verrons qu'il peut être explicitement traité dans deux cas. Le premier correspond à une onde incidente qui n'est résonnante avec aucune des fréquences propres (des fréquences de Bohr) de l'atome. Comme nous ne pourrons pas traiter l'émission spontanée, nous ne décrirons pas correctement l'amortissement du mouvement électronique et nous nous heurterons à des divergences si la fréquence d'excitation s'approche d'une résonance atomique. L'autre cas que nous pourrons explicitement traiter correspond au cas d'une onde incidente strictement résonante avec une transition atomique, à condition qu'on puisse alors négliger l'influence de toutes les autres fréquences propres. Nous montrerons alors qu'on peut éliminer les divergences et calculer explicitement l'évolution atomique.

En traitant la diffusion non résonnante par un atome quantique, nous nous apercevrons que les

résultats sont en fait essentiellement identiques à ceux du modèle naïf de l'électron élastiquement lié. Nous donnerons ainsi à ce modèle une validité beaucoup plus solide que celle que nous aurions pu lui attribuer de prime abord. En fait, dans toute la suite de ce cours, quand nous aurons par exemple à traiter de la propagation dans un milieu matériel, c'est ce modèle que nous utiliserons.

Nous appliquerons enfin, dans le troisième paragraphe, ces résultats au problème de la diffusion par un milieu comprenant un grand nombre d'atomes. Ces atomes étant décrits à ce stade par leur seule polarisabilité, les résultats de ce paragraphe s'appliqueront en fait à n'importe quel modèle de centres diffuseurs (atome classique, atome semi-classique, électrons libres...). Nous tenterons en particulier de comprendre pourquoi les solides transparents diffusent beaucoup moins que les gaz, alors qu'ils sont notablement plus denses. Nous montrerons que le champ diffusé est simplement relié à la transformée de Fourier des variations de densité du milieu. Nous pourrons ainsi examiner la diffusion par un milieu homogène dense, par un cristal avec une séparation entre atomes de l'ordre de la longueur d'onde, et enfin la diffusion par un milieu désordonné comme un gaz.

3.1 Modèle de Thomson

3.1.1 Modèle de l'électron élastiquement lié.

Nous considérerons donc un modèle extrêmement simpliste d'atome: un seul électron, lié par un force harmonique isotrope. Ce modèle a connu une certaine faveur au tournant du siècle (on consultera pour une approche historique plus détaillée l'appendice sur le modèle de Bohr dans la première partie de ce cours). Il était assez naturel d'essayer de comprendre les fréquences discrètes rayonnées ou absorbées par les atomes en termes de résonances d'oscillateurs harmoniques. Les expériences de Rutherford n'existaient pas encore. Il n'y avait donc aucune raison de postuler une structure planétaire pour l'atome (nous avons montré de plus qu'un modèle planétaire serait nécessairement instable). En revanche, des coïncidences (très imparfaites et accidentelles) entre les raies de l'hydrogène et les harmoniques élevés d'une fréquence fondamentale pouvaient suggérer que les raies observées étaient les harmoniques de la fréquence de résonance d'un oscillateur.

Le modèle de Thomson[1] (dit encore modèle du "plum–pudding") s'appuyait sur cette interprétation. On considère l'atome comme constitué d'une "gelée" positive, uniformément chargée, portant une charge totale unité. Dans cette "gelée", un électron unique se déplace librement. Le champ électrique créé par une sphère uniformément chargée étant, à l'intérieur de la sphère, linéaire en fonction de la distance au centre, l'électron se trouve lié par une force harmonique (force centrale proportionnelle à la distance). Nous laisserons au lecteur le soin de calculer la constante de force et la fréquence d'oscillation en fonction du rayon de la sphère. En prenant un rayon de la sphère de l'ordre de l'Ångström, on trouve effectivement une fréquence de résonance dans le domaine optique. On peut montrer aussi que l'ordre de grandeur de l'énergie nécessaire pour arracher l'électron est de l'ordre des énergies d'ionisation typiques. Enfin, on peut utiliser ce modèle pour calculer, comme nous allons le faire à différents endroits de ce cours, les indices de réfraction ou susceptibilités diélectriques typiques. Nous verrons que, là encore, au moins en ordres de grandeur, ce modèle était raisonnable. Il a en fait tenu jusqu'à la révélation d'un noyau pratiquement ponctuel dans la structure atomique. Le modèle de la gelée positive devenait alors intenable. Le modèle planétaire était le seul possible. Sa stabilité ne pouvant être comprise dans le cadre de la mécanique classique, il a suscité alors une crise scientifique majeure qui ne fut totalement résolue qu'en 1926 avec la découverte de la mécanique ondulatoire par Schrödinger et de la mécanique des matrices par Heisenberg.

Nous ne nous occuperons pas des détails du modèle de Thomson dans ce cours. Nous considérerons seulement un électron de masse m et de charge $q = -e$ ($e > 0$) harmoniquement lié à un centre fixe avec une fréquence propre d'oscillation ω_0. Avec ces conditions, l'équation de mouvement de l'électron

[1] Thomson est connu comme découvreur de l'électron, en 1897. En fait, Wiechert (celui des potentiels de Liénard), a joué un rôle au moins aussi important dans la découverte de la nature corpusculaire des rayons cathodiques.

s'écrit alors:

$$\frac{d^2\mathbf{r}}{dt^2} + \omega_0^2\mathbf{r} = 0 \ , \tag{3.1}$$

$\mathbf{r}$ étant la position de l'électron par rapport au centre de forces. A cette équation, il conviendra d'ajouter la contribution de la force de réaction de rayonnement. En première approximation, la solution de cette équation est bien sûr une évolution harmonique de $\mathbf{r}$ selon les trois axes. La trajectoire de l'électron est dans le cas le plus général une ellipse centrée à l'origine. "L'état fondamental" de ce système correspond bien sûr à un électron immobile à l'origine.

3.1.2 Emission spontanée

Nous tenterons dans ce paragraphe de modéliser l'émission spontanée. Nous supposerons que, à l'instant origine, l'électron initialement au repos est mis en mouvement. Dans une lampe à décharge, ce mouvement serait dû à une collision avec un des électrons libres de la décharge. Nous supposerons, pour simplifier les calculs sans beaucoup restreindre la généralité, que le mouvement de l'électron est ensuite linéaire, orienté de long de l'axe Oz[2]. Cet électron oscillant se comporte comme un dipôle. Les ordres de grandeur du modèle montrent en effet que l'extension du mouvement (au plus de l'ordre du diamètre de la sphère chargée, c'est à dire de l'ordre de l'Ångström) est très petite devant la longueur d'onde rayonnée (de l'ordre du micron). Il est alors aisé de calculer, en utilisant les résultats du premier chapitre, la puissance rayonnée par l'électron:

$$\mathcal{P} = \frac{q^2 z_0^2 \omega_0^4}{12\pi\epsilon_0 c^3} \ , \tag{3.2}$$

où z_0 est l'amplitude du mouvement. Dans notre modèle, l'énergie rayonnée ne peut provenir que de l'énergie mécanique de l'électron. Cette énergie est donc une fonction décroissante du temps: le mouvement de l'électron doit s'amortir.

L'amortissement du mouvement de l'électron, lent devant la période, peut être décrit au moyen de la fore de réaction de rayonnement $f = m\tau\dot{a}$. En incluant cette force, l'équation du mouvement de l'électron s'écrit:

$$\frac{d^2\mathbf{r}}{dt^2} - \tau\frac{d^3\mathbf{r}}{dt^3} + \omega_0^2\mathbf{r} = 0 \ , \tag{3.3}$$

où τ est le temps caractéristique de la réaction de rayonnement défini plus haut. On peut aisément trouver les fréquences propres de cette équation, sous forme d'un développement en termes de $\eta = \omega_0\tau$. Ce paramètre est effectivement, pour une transition optique ordinaire, de l'ordre de 10^{-10} (on rappelle que l'ordre de grandeur de τ est 10^{-24} s). En portant dans l'équation du mouvement une solution en $\mathbf{r} = \mathbf{r}_0 \exp(-i\omega t)$ et en posant:

$$\omega = \omega_0(1 + x) \ , \tag{3.4}$$

on met l'équation caractéristique sous la forme:

$$-(1 + x)^2 - i\eta(1 + x)^3 + 1 = 0 \ . \tag{3.5}$$

Au premier ordre non nul, on trouve donc

$$x = -i\frac{\eta}{2} \ . \tag{3.6}$$

On peut alors écrire

$$\omega = \omega_0(1 - i\frac{\eta}{2} + y) \ , \tag{3.7}$$

[2]On pourra, à titre d'exercice, considérer le cas général en superposant des mouvements d'amplitudes et de phases arbitraires le long des trois axes.

où y est du second ordre en η. En reportant cette forme dans l'équation caractéristique et en identifiant les termes d'ordre 2, on trouve:

$$y = -\frac{5\eta^2}{8} \ . \tag{3.8}$$

La seule solution physiquement acceptable à l'ordre 2 en η peut donc s'écrire:

$$\omega = \omega_0 + \Delta\omega - i\gamma/2 \ . \tag{3.9}$$

Le terme

$$\Delta = -\frac{5}{8}\eta^2\omega_0 \tag{3.10}$$

correspond à un déplacement de fréquence. Il est, comme nous le verrons, tout à fait négligeable. Le terme imaginaire correspond, comme nous l'attendions, à un amortissement du mouvement. L'énergie mécanique, proportionnelle au carré de l'amplitude du mouvement, est amortie exponentiellement avec un taux γ (et une constante de temps $T = 1/\gamma$) égal à:

$$\gamma = \omega_0\eta = \omega_0^2\tau = \frac{q^2\omega_0^2}{6\pi\epsilon_0 cmc^2} \ . \tag{3.11}$$

Une manifestation expérimentale de cette décroissance exponentielle est que le spectre de fluorescence d'un ensemble d'atomes ait une forme lorentzienne. Le spectre étant relié à la transformée de Fourier de la réponse temporelle[3], celle-ci, transformée de Fourier d'une Lorentzienne, est une exponentielle. On peut bien sûr maintenant accéder aussi, directement, à la réponse temporelle avec une électronique modérément rapide. Un des grands mérites de ce modèle très simplifié est donc de prédire correctement l'émission de lumière par des atomes excités ou émission spontanée.

Estimons l'ordre de grandeur de la durée de vie radiative T et du taux d'émission spontanée γ (notons que nous avons déjà effectué un calcul semblable dans les paragraphes sur la réaction de rayonnement avec un modèle différent d'orbites circulaire. Sans surprises, les ordres de grandeur relatifs aux deux modèles sont similaires). Pour cela, nous calculerons le rapport γ/ω_0, l'inverse du "facteur de qualité" de l'oscillateur harmonique. Ce rapport est proportionnel à la fréquence du rayonnement ω_0. Pour un ordre de grandeur, nous allons utiliser les fréquences prédites par le modèle de Bohr, décrit en détails dans l'appendice de la première partie de ce cours. Assez bizarrement, nous mêlons ainsi des arguments du modèle planétaire de la "première théorie quantique" de 1913 avec le modèle, incompatible, de l'électron élastiquement lié. Rien, a priori, ne fixant la fréquence de résonance dans le modèle de Thomson, il nous faut bien utiliser les valeurs expérimentales, qui sont correctement prédites par le modèle de Bohr. La fréquence est donc de l'ordre de la constante de Rydberg divisée par la constante de Planck, $R/\hbar$. La constante de Rydberg, elle même, est de l'ordre de:

$$R = \frac{1}{2}mc^2\alpha^2 \ , \tag{3.12}$$

produit, à un facteur deux près (dont l'origine peut être reliée au théorème du viriel), de l'énergie de masse de l'électron par le carré de la "constante de structure fine" α définie par:

$$\alpha = \frac{1}{4\pi\epsilon_0}\frac{q^2}{\hbar c} \ . \tag{3.13}$$

En portant cette expression de l'ordre de grandeur de la fréquence atomique dans le facteur de qualité, nous trouvons:

$$\frac{\gamma}{\omega_0} = \omega_0\tau = \frac{R\tau}{\hbar} = \frac{q^2mc^2\alpha^2}{12\pi\epsilon_0 c\hbar mc^2} = \frac{\alpha^3}{3} \ . \tag{3.14}$$

[3]En fait le spectre est la transformée de Fourier de la fonction de corrélation du champ émis, qui se rapporte à la réponse temporelle pour une excitation impulsionnelle.

Nous trouvons ainsi un résultat très simple (mais certainement pas inattendu): le facteur de qualité typique d'un oscillateur atomique est de l'ordre de l'inverse du cube de la constante de structure fine. Numériquement, il est donc de l'ordre de 10^7. La fréquence d'une transition atomique étant, dans le domaine visible, de l'ordre de 10^{15} Hz, le taux d'émission spontanée est de l'ordre de 10^8 s^{-1}. La durée de vie radiative d'un atome excité est donc de l'ordre de la dizaine de nanosecondes. La valeur expérimentale pour la fameuse raie jaune du sodium (en fait un doublet de structure fine à 589 et 589.6 nm) est de 16 ns, correspondant à une largeur de la raie de l'ordre de 10 MHz. Notre modèle, de façon un peu surprenante, est plus que raisonnable[4].

Estimons maintenant le déplacement de fréquence $\Delta\omega$. Il est facile de voir que:

$$\Delta\omega = -\frac{5}{8}\eta\gamma \ . \tag{3.15}$$

Il ne vaut qu'une fraction de Hertz et il est donc complètement négligeable. Notons que, dans un modèle quantique, il existe des déplacements des niveaux beaucoup plus importants (de l'ordre du GigaHertz) dûs au couplage de l'atome aux modes vides du champ électromagnétique quantifié (déplacements de Lamb).

Avant de poursuivre, nous pouvons simplifier un peu l'équation du mouvement de l'électron pour nous affranchir du terme du troisième ordre, $-\tau d^3\mathbf{r}/dt^3$. Comme le facteur de qualité de la transition atomique est très grand, le mouvement est très proche d'un mouvement harmonique. Avec une très bonne approximation, ce terme peut donc s'écrire $\tau\omega_0^2\dot{\mathbf{r}}$. La force de réaction de rayonnement peut donc être considérée comme proportionnelle à la vitesse. On écrira donc l'équation du mouvement tenant compte de la réaction de rayonnement sous la forme:

$$\frac{d^2\mathbf{r}}{dt^2} + \gamma\frac{d\mathbf{r}}{dt} + \omega_0^2\mathbf{r} = 0 \ . \tag{3.16}$$

Nous pouvons maintenant, en utilisant cette équation, traiter le problème de la diffusion d'un champ, éventuellement résonnant.

3.1.3 Diffusion du rayonnement.

Nous étudierons dans ce paragraphe le rayonnement d'un atome classique, excité de façon permanente par une onde plane monochromatique. Sous l'influence de cette onde, l'atome acquiert un dipôle induit qui rayonne. A la différence du paragraphe précédent, il s'agit alors d'un rayonnement en régime stationnaire. De plus, l'onde incidente n'est pas nécessairement résonnante à la fréquence propre de l'atome.

Dipôle induit. Polarisabilité.

Nous supposerons l'onde incidente polarisée linéairement selon Oz (la généralisation ne présente aucune difficulté). Nous ne considérerons pas le cas d'une onde de très haute fréquence, pour lequel la taille de l'atome pourrait être comparable à la longueur d'onde. Nous pourrons donc supposer l'atome ponctuel et ignorer la structure spatiale de l'onde. Nous écrirons donc simplement le champ électrique incident sous la forme $E_0\mathbf{u}_z\exp(-i\omega t)$. La fréquence ω est arbitraire, la phase de l'onde choisie de telle manière que E_0 soit réel (ce qui ne restreint en rien la généralité). L'équation du mouvement de l'électron en présence de cette onde s'écrit, avec le terme de réaction de rayonnement:

$$\frac{d^2\mathbf{r}}{dt^2} + \gamma\frac{d\mathbf{r}}{dt} + \omega_0^2\mathbf{r} = \frac{qE_0}{m}\mathbf{u}_z e^{-i\omega t} \ . \tag{3.17}$$

[4] Il ne permet cependant pas de comprendre pourquoi il existe des raies atomiques beaucoup plus fines, correspondant à des durées de vie infiniment plus longues que ce que nous venons de calculer. La raie de 1S vers 2S de l'hydrogène, par exemple, aurait dans notre modèle une durée de vie de l'ordre de la nanoseconde (c'est une raie dans l'ultraviolet lointain). Sa durée de vie est, en fait, de 1/7 de seconde!

La solution contient un terme transitoire, dépendant de la condition initiale pour l'atome, qui s'amortit en un temps de l'ordre de τ. On peut en général ignorer ce transitoire et ne considérer que la solution en régime permanent. Elle s'écrit évidemment sous la forme $\mathbf{r}_0 \exp(-i\omega t)$ avec:

$$\mathbf{r}_0 = \frac{qE_0/m}{\omega_0^2 - \omega^2 - i\gamma\omega}\mathbf{u}_z \ . \tag{3.18}$$

L'amplitude complexe du dipôle électrique acquis par l'atome à la fréquence ω s'écrit donc:

$$\mathbf{d}_0 = \frac{q^2/m}{\omega_0^2 - \omega^2 - i\gamma\omega}\mathbf{E}_0 \ . \tag{3.19}$$

Elle est proportionnelle à l'amplitude du champ incident. On peut donc définir la polarisabilité **classique** de l'atome à la fréquence ω, $\alpha_c(\omega)$, par:

$$\mathbf{d}_0 = \epsilon_0 \alpha_c(\omega)\mathbf{E}_0 \ . \tag{3.20}$$

Cette polarisabilité est simplement donnée par:

$$\alpha_c(\omega) = \frac{q^2}{m\epsilon_0}\frac{1}{\omega_0^2 - \omega^2 - i\gamma\omega} \ . \tag{3.21}$$

On notera qu'avec cette définition, la polarisabilité est homogène à un volume (nous verrons dans la partie sur l'électromagnétisme dans les milieux matériels l'intérêt de ce choix). A fréquence nulle, on a simplement $\alpha_c(0) = q^2/m\epsilon_0\omega_0^2$. On vérifiera par un calcul évident d'électrostatique que la polarisabilité à fréquence nulle coïncide avec le volume de la sphère contenant la "gelée" positive. On remarquera aussi que, en faisant $\omega_0 = 0$, on trouve bien la polarisabilité d'un électron libre. Ce modèle nous permet aussi de traiter la diffusion de rayonnement par un plasma dont les charges peuvent, en première approximation, être considérées comme libres. Notons finalement que l'expression de la polarisabilité peut se simplifier, au voisinage de la résonance, quand l'amortissement γ est très petit par rapport à la fréquence propre. On a alors:

$$\alpha_c(\omega) \simeq \frac{q^2}{m\epsilon_0\omega_0}\frac{1}{2(\omega_0 - \omega) - i\gamma} \ . \tag{3.22}$$

Le module au carré de la polarisabilité est alors une simple lorentzienne, de centre ω_0 et de largeur γ. Le facteur de qualité des résonances atomiques étant toujours excellent, cette approximation est très légitime et toujours utilisée.

A partir de l'expression du dipôle induit, nous pouvons aisément estimer la puissance rayonnée. Elle fera intervenir le carré du module de la polarisabilité, sous la forme:

$$\mathcal{P} = \frac{\epsilon_0}{12\pi c^3}|\alpha_c|^2\omega^4 E_0^2 \ . \tag{3.23}$$

Cette puissance est évidemment proportionnelle au carré de l'amplitude du champ incident. Pour obtenir un résultat indépendant de l'intensité de l'onde incidente, nous allons rapporter la puissance diffusée à la puissance incidente par unité de surface, $\epsilon_0 c E_0^2/2$. Le rapport d'une puissance à une puissance par unité de surface est une surface. Nous caractérisons ainsi l'efficacité de diffusion de l'atome par une section efficace σ. La puissance diffusée est simplement la puissance incidente sur une surface σ (normale à la direction de propagation). En un mot, tous les photons qui "tombent" dans une surface σ centrée sur l'atome sont diffusés. On a:

$$\sigma = \frac{1}{6\pi}\left(\frac{\omega}{c}\right)^4 |\alpha_c|^2 \ . \tag{3.24}$$

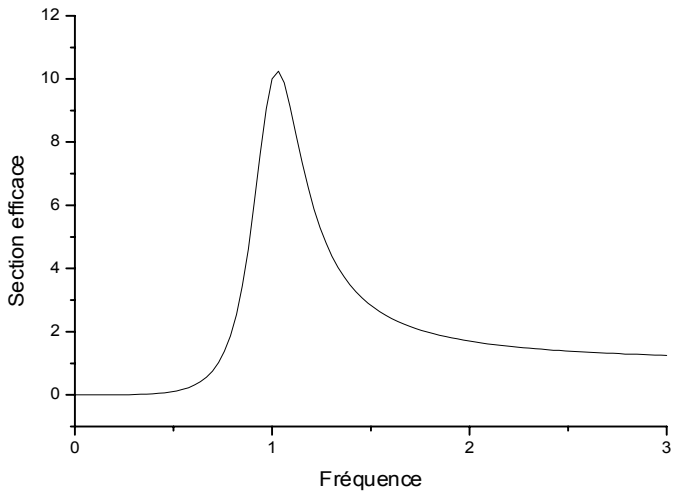

Figure 3.1: Section efficace de diffusion en fonction de la fréquence, exprimée en unités de ω_0. Le facteur de qualité à été limité à 3.3 pour rendre la figure lisible.

On peut encore écrire, en explicitant la polarisabilité:

$$\sigma = \frac{1}{6\pi c^4}\frac{q^4}{m^2\epsilon_0^2}\frac{\omega^4}{(\omega_0^2-\omega^2)^2+\gamma^2\omega^2} \ . \tag{3.25}$$

Pour simplifier quelque peu le préfacteur, nous allons faire intervenir une dimension caractéristique de l'électron, le "rayon classique de l'électron" r_e que nous avons déjà rencontré à propos de la réaction de rayonnement. Un simple argument d'analyse dimensionnelle montre que la seule longueur qu'on peut former à partir des caractéristiques de l'électron est:

$$r_e = \frac{q^2}{4\pi\epsilon_0 mc^2} = \frac{3}{2}c\tau \ , \tag{3.26}$$

qui est de l'ordre de 3 fm. On peut donc écrire enfin:

$$\sigma = \frac{8\pi}{3}r_e^2\frac{\omega^4}{(\omega_0^2-\omega^2)^2+\gamma^2\omega^2} \ . \tag{3.27}$$

La figure 3.1 présente la section efficace de diffusion en fonction de la fréquence.

Pour des fréquences proches de résonance, nous pouvons écrire, comme pour la polarisabilité:

$$\sigma = \frac{8\pi}{3}r_e^2\frac{\omega^2}{4(\omega_0-\omega)^2+\gamma^2} \ . \tag{3.28}$$

La section efficace est le produit de la "surface de l'électron" par un facteur présentant une résonance Lorentzienne en ω_0. La section efficace sera donc qualitativement très différente selon que la fréquence incidente est très grande, très petite ou voisine de la fréquence propre. Nous allons examiner séparément ces différents cas.

Diffusion Rayleigh

Nous considérerons d'abord le cas de la diffusion très basse fréquence: $\omega \ll \omega_0$. C'est par exemple pratiquement toujours le cas pour la diffusion radiofréquence ou infrarouge lointain. C'est aussi le

cas de la diffusion de la lumière visible par l'air. Les première fréquences de résonance optique de l'oxygène ou de l'azote sont en effet situées dans le domaine ultraviolet. En tenant compte de cette condition, la section efficace s'écrit simplement:

$$\sigma = \frac{8\pi}{3} r_e^2 \left(\frac{\omega}{\omega_0}\right)^4 . \tag{3.29}$$

Cette expression appelle plusieurs commentaires. D'abord, la valeur de la section efficace est extrêmement faible. Même si ω est de l'ordre de ω_0, la section efficace n'est que de l'ordre de la "surface classique" de l'électron, 10^{-30} m^2. Pour mieux comprendre la faiblesse de cette valeur, calculons la longueur de propagation nécessaire pour que la lumière incidente soit notablement atténuée par la diffusion, L. Avec un faisceau de section S, l'atténuation sera complète si $S = \mathcal{N}\sigma$ où $\mathcal{N} = NSL$ est le nombre total de molécules dans le faisceau sur une longueur L (N est la densité numérique). On trouve donc simplement:

$$L = \frac{1}{\sigma N} . \tag{3.30}$$

Pour un gaz à la pression atmosphérique, la densité typique est de 3.10^{25} m^{-3}. En prenant une valeur très surévaluée de 10^{-30} m^2 pour σ, on trouve une longueur d'atténuation L de 30 km. On comprend donc pourquoi, dans la journée, la lumière solaire, qui ne traverse que quelques kilomètres d'atmosphère dense, n'est pas très atténuée par diffusion (la lumière directe est très intense par rapport au "bleu" du ciel). En revanche, au lever ou au coucher du soleil, la lumière, sous incidence rasante, traverse une épaisseur de gaz beaucoup plus considérable, pouvant dépasser la centaine de kilomètres. Dans ce cas la lumière est fortement atténuée.

Remarquons aussi la très rapide dépendance de la section efficace en fonction de la fréquence incidente. Les courtes longueurs d'onde, plus proches de la résonance, sont diffusées beaucoup plus efficacement que les plus longues. Voici une explication simple de la couleur bleue de la lumière diffusée par les gaz ou les fumées. La lumière transmise, elle, apparaît plus rouge, les fréquences les plus élevées étant aussi les plus atténuées.

Notons que ces raisonnements nous donnent quelques indications sur l'efficacité de la diffusion ou la nature du spectre diffusé par un gaz. Ils ne nous prouvent pas, en revanche, qu'un échantillon macroscopique de gaz diffuse. La manière dont s'additionnent les rayonnements produits par les différentes molécules du gaz n'est pas du tout prise en compte dans ce raisonnement simpliste qui ne considère qu'une molécule unique. On pourrait se demander en particulier pourquoi un gaz diffuse alors qu'un verre, pourtant beaucoup plus dense, ne diffuse pratiquement pas s'il est pur. Ainsi, la longueur d'atténuation dans les fibres optiques est de l'ordre de la dizaine de kilomètre aussi, alors que la densité numérique est trois ordres de grandeur plus élevée que celle de l'air. Pour bien comprendre cet effet, il faut analyser en détails le processus de diffusion macroscopique. On montre alors que les fluctuations thermodynamiques de densité jouent un rôle essentiel dans la diffusion. Pour un milieu statique, comme un verre, les amplitudes diffusées interfèrent destructivement et il ne reste que la lumière transmise vers l'avant. Pour un gaz, le nombre de particules dans un volume de l'ordre de λ^3 est fluctuant. Ces fluctuations sont responsables de la présence de lumière diffusée. Il faut donc prendre avec précautions les ordres de grandeur que nous venons de donner.

Diffusion Thomson

Nous considérerons maintenant le domaine des hautes fréquences $\omega \gg \omega_0$. C'est, par exemple, celui de la diffusion des rayonnements X dans les matériaux non absorbants. On trouve alors simplement:

$$\sigma = \frac{8\pi}{3} r_e^2 . \tag{3.31}$$

La section efficace est constante, indépendante de la fréquence incidente, simplement égale à la surface classique de l'électron. En fait, pour de hautes fréquences incidentes, le fait que l'électron soit lié par

une force harmonique n'a pratiquement aucune influence sur la diffusion. On trouve ici simplement la section efficace de diffusion par un électron libre. Les ordres de grandeur de la section efficace et des longueurs d'atténuation sont essentiellement les mêmes que dans le paragraphe précédent.

Diffusion résonnante

Le régime le plus intéressant est celui où la fréquence incidente est proche de la résonance atomique $\omega \simeq \omega_0$. La section efficace présente alors un comportement quasi Lorentzien autour de ω_0 décrit par l'équation (3.28).

Pour étudier les ordres de grandeur, nous nous focaliserons sur le cas où le rayonnement incident est strictement résonant: $\omega = \omega_0$. Dans ce cas,

$$\sigma = \frac{8\pi}{3} r_e^2 \left(\frac{\omega_0}{\gamma}\right)^2 . \tag{3.32}$$

La section efficace est beaucoup plus grande que la surface classique de l'électron. Nous avons vu en effet que le facteur de qualité de la transition atomique est de l'ordre de 10^7. La section efficace résonante est donc 14 ordres de grandeur plus élevée que les sections Thomson ou Rayleigh. Plus précisément:

$$\frac{\omega_0}{\gamma} \simeq \frac{1}{\omega_0 \tau} = \frac{6\pi\epsilon_0 m c^3}{q^2 \omega_0} , \tag{3.33}$$

et donc

$$r_e \frac{\omega_0}{\gamma} = \frac{r_e}{\omega_0 \tau} = \frac{3}{2}\frac{c}{\omega_0} . \tag{3.34}$$

On en déduit finalement

$$\sigma = \frac{3}{2\pi}\lambda_0^2 , \tag{3.35}$$

où λ_0 est la longueur d'onde de la lumière résonante. La section efficace de diffusion résonante est donc de l'ordre du carré de la longueur d'onde, 1 micron carré environ. En considérant par exemple une vapeur de sodium sous une pression de 10^{-5} torr (ce qu'on obtient dans une ampoule chauffée à une centaine de degrés) soit une densité numérique de 10^{17} m^{-3}, on trouverait une longueur d'absorption (le raisonnement des paragraphes précédents reste correct) de 100 μm. Il s'agit bien sûr d'une valeur trop faible. On constate expérimentalement que le rayonnement est atténué sur une longueur courte, mais notablement plus longue que ce que nous venons de calculer, surtout quand le rayonnement incident est intense.

L'origine de ce désaccord vient du fait que nous travaillons avec un modèle linéaire qui ignore toute saturation du système atomique. Aussi intense et aussi résonante que soit l'onde incidente, l'atome répond toujours linéairement, avec une section efficace indépendante de l'intensité incidente. Si on quantifie correctement la dynamique atomique, on s'aperçoit que ce comportement n'est valable que pour des diffusions non résonantes (c'est ce que nous établirons dans le prochain paragraphe). Pour la diffusion résonante, l'atome ne peut diffuser plus d'un photon dans un intervalle de temps γ^{-1}. Comme nous le verrons dans le prochain paragraphe, il peut osciller rapidement entre les deux niveaux d'énergie bordant la transition résonnante mais ces cycles correspondent à des cycles absorption/émission induite qui ne changent pas le nombre de photons de l'onde incidente. Les seuls événements qui correspondent à une diffusion sont ceux se produisant par émission spontanée, tous les γ^{-1}. L'ordre de grandeur de la section efficace que nous donnons ici ne sera correct que si l'atome diffuse beaucoup moins d'un photon dans une durée de vie radiative. Il faut donc que la puissance incidente soit petite devant un photon (avec une énergie d'environ un eV) par 10 ns et par micron carré. L'ordre de grandeur de la "puissance de saturation" est donc de 10 W/m^2 ou encore de 1 mW/cm^2. Une approche quantique rigoureuse donne exactement cet ordre de grandeur. Le modèle de Thomson fait encore la preuve de son efficacité.

Notons aussi que nous avons supposé dans tout ce raisonnement que la largeur de la résonance était entièrement déterminée par la durée de vie spontanée. L'analyse est un peu plus complexe si on tient compte d'une autre cause d'élargissement, comme l'effet Doppler. On trouve alors en général une réduction importante de la section efficace de diffusion et, en conséquence, une augmentation de la puissance de saturation.

Nous avons maintenant déduit tout ce qui était accessible avec le modèle de Thomson. Nous allons maintenant nous consacrer à un modèle semi–quantique (ou semi–classique) qui nous permettra de donner une assise plus solide aux résultats que nous venons d'établir.

3.2 Modèle semi–classique

Nous traiterons dans ce paragraphe le rayonnement d'une source atomique dans un modèle où nous quantifierons la structure atomique. En revanche, le champ électromagnétique sera encore considéré comme une quantité classique. Cette quantification partielle nous empêchera de décrire le phénomène d'émission spontanée. Celle-ci ne peut être comprise quantitativement que dans un modèle où le champ est convenablement quantifié (l'émission spontanée est, en partie, "induite" par les fluctuations de point zéro du champ électromagnétique, les célèbres "fluctuations du vide", authentiquement quantiques). Pour éviter ces difficultés, nous ne traiterons dans ce paragraphe que de la diffusion de rayonnement par un atome quantique. En nous plaçant dans des régimes limites convenables, nous pourrons éviter d'avoir à tenir compte de l'émission spontanée. Dans un premier paragraphe, nous nous poserons le problème de relier les observables atomiques au rayonnement diffusé. Quel dipôle devons nous insérer dans les formules classiques du rayonnement pour estimer la puissance diffusée par un atome? Ensuite, nous écrirons les équations d'évolution quantiques d'un atome placé dans un champ incident. Nous traiterons ensuite deux cas limites: la diffusion non résonnante et la diffusion strictement résonnante.

3.2.1 Rayonnement d'un atome quantique

Nous considérerons un modèle atomique très simple, celui d'un atome à un électron, l'hydrogène. Les niveaux atomiques résultent de la quantification de mouvement de l'électron dans le potentiel central du noyau[5]. On trouvera cette quantification détaillée dans les manuels de mécanique quantique (Cohen en particulier). Nous nous contenterons ici de savoir qu'il existe des niveaux stationnaires $|i\rangle$ d'énergie E_i, états propres du Hamiltonien atomique H_0 (énergie cinétique plus énergie potentielle de liaison). Le fondamental sera appelé $|g\rangle$. Son énergie est nulle par convention. Ces niveaux, que nous traiterons comme s'ils étaient non dégénérés, sont des états propres de l'opérateur parité, qui commute avec le Hamiltonien. Les fonctions d'onde leur correspondant sont donc paires ou impaires par rapport à l'origine (où est situé le noyau). Pour l'hydrogène, ainsi, le niveau fondamental est le $1S$, avec une fonction d'onde à symétrie sphérique, donc paire. Les premiers niveaux excités sont le $2S$, également pair et le $2P$, impair (la fonction d'onde est, en coordonnées sphériques, proportionnelle à $\cos\theta$—voir l'appendice de la première partie).

La position de l'électron dans l'atome est un opérateur $\mathbf{R}$. Le dipôle électrique est donc, lui aussi, un opérateur relatif à l'électron, $\mathbf{D} = q\mathbf{R}$ (q étant la charge de l'électron). Pour estimer la puissance rayonnée par l'atome, nous avons besoin, dans ce modèle semi–classique, d'une amplitude de dipôle classique. Le plus naturel est d'employer la valeur moyenne de l'opérateur dipolaire dans l'état atomique. On décrira bien ainsi le rayonnement d'un grand nombre d'atomes soumis au même champ incident, ce qui est bien, en général, le problème de la diffusion. Nous emploierons donc les formules standard du rayonnement dipolaire avec $\mathbf{d} = \langle\mathbf{D}\rangle$. Il est évident que nous ne pourrons ainsi traiter que le rayonnement moyen d'une grande assemblée d'atomes (les valeurs moyennes de la mécanique quantique décrivent des moyennes sur un grand nombre de réalisations d'une expérience unique, les

[5]On négligera ici tous les effets dûs au spin de l'électron ou du noyau.

moyennes d'ensemble). Parler, dans ce modèle, du rayonnement d'un atome unique n'aurait aucun sens. En fait, un atome unique émet un photon en effectuant un saut quantique d'un niveau à un autre. Un tel processus, impliquant un photon unique et tenant nécessairement en compte l'appareil détectant ce photon, n'a pas de sens dans une description classique du champ.

Nous pouvons tout de suite nous poser le problème de savoir quel type d'états atomiques rayonnent. Il faut que la valeur moyenne de l'opérateur dipolaire

$$\langle \mathbf{D} \rangle = \langle \Psi | \mathbf{D} | \Psi \rangle = q \int |\Psi(\mathbf{r})|^2 \mathbf{r}\, d^3\mathbf{r} \tag{3.36}$$

($|\Psi\rangle$ est la fonction d'onde atomique) soit non nulle. Si $|\Psi\rangle$ a une parité bien définie, le carré de son module est toujours pair et l'intégrale identiquement nulle. Les états propres du Hamiltonien atomique étant de parité bien définie, le dipôle moyen dans tous ces états est nul[6]. Les niveaux stationnaires ne rayonnent pas et sont donc stables. On constate là l'insuffisance évidente d'un modèle traitant le champ classiquement. Nous savons bien que de tous les niveaux atomiques seul le fondamental, $|g\rangle$, est stable. Tous les autres perdent leur énergie par rayonnement, même si la durée de vie est très longue (des années pour les niveaux excités hyperfins). Il est tout simplement impossible de traiter l'émission spontanée dans un cadre classique. D'un point de vue plus positif, l'émission spontanée est la preuve la plus formelle de la nature quantique du champ électromagnétique.

Seules rayonnent donc des fonctions d'onde dont la parité n'est pas bien définie. Il s'agit de combinaisons linéaires de niveaux de parité opposée. Si , à un instant donné,

$$|\Psi\rangle = \frac{1}{\sqrt{2}}(|i\rangle + |j\rangle) \ , \tag{3.37}$$

alors

$$\langle \mathbf{D} \rangle = \mathrm{Re}\, \langle i | \mathbf{D} | j \rangle = \mathrm{Re}\, q \int \Psi_i^* \Psi_j \mathbf{r}\, d\mathbf{r} \ , \tag{3.38}$$

qui peut être non nul si $|i\rangle$ et $|j\rangle$ sont de parité opposée. Ces niveaux n'ayant pas la même énergie, la superposition évolue temporellement comme:

$$|\Psi\rangle = \frac{1}{\sqrt{2}}(|i\rangle e^{-iE_i t/\hbar} + |j\rangle e^{-iE_i j t/\hbar}) \ . \tag{3.39}$$

Le dipôle moyen évolue donc comme:

$$\langle \mathbf{D} \rangle = \mathrm{Re}\, \langle i | \mathbf{D} | j \rangle e^{-i\omega_{ji} t} \ , \tag{3.40}$$

où $\omega_{ji} = (E_j - E_i)/\hbar$ est la fréquence de Bohr entre ces deux niveaux. L'atome dans une superposition de ces deux niveaux rayonne donc à la fréquence de Bohr, c'est à dire à la fréquence de la transition atomique, ce qui n'est pas très surprenant. Notons qu'une superposition plus complexe, mettant en jeu plusieurs niveaux de parités variées, correspondra à un dipôle oscillant sur plusieurs composantes de fréquence. Celles-ci sont les fréquences de Bohr des transitions "autorisées" entre ces niveaux, celles pour lesquelles l'élément de matrice dipolaire $\langle i | \mathbf{D} | j \rangle$ est non nul.

Pour illustrer visuellement le rayonnement d'une superposition d'états, la figure 3.2 présente le module de la fonction d'onde d'un atome d'hydrogène préparé dans une superposition quantique des états $|1S\rangle$ et $|2P\rangle$ ($m = 0$ correspondant à une orbitale antisymétrique par rapport au plan xOy) à deux instants séparés par une demi–période de Bohr. En n'oubliant pas que la densité de charge électronique est proportionnelle au carré du module de la fonction d'onde, on voit bien que la charge

[6] La nullité du dipôle moyen d'un état stationnaire est liée à l'invariance par renversement du sens du temps de la dynamique atomique. Si la symétrie est respectée, le dipôle doit être nul. De nombreuses expériences ont été consacrées à la recherche d'un moment électrique dipolaire pour un atome isolé (ou d'un neutron, les arguments étant à peu près les mêmes), dans le but de mettre en évidence une éventuelle violation de la symétrie temporelle. Elles ont permis de mettre des limites supérieures très basses aux dipôles, mais pas encore de mesurer un effet.

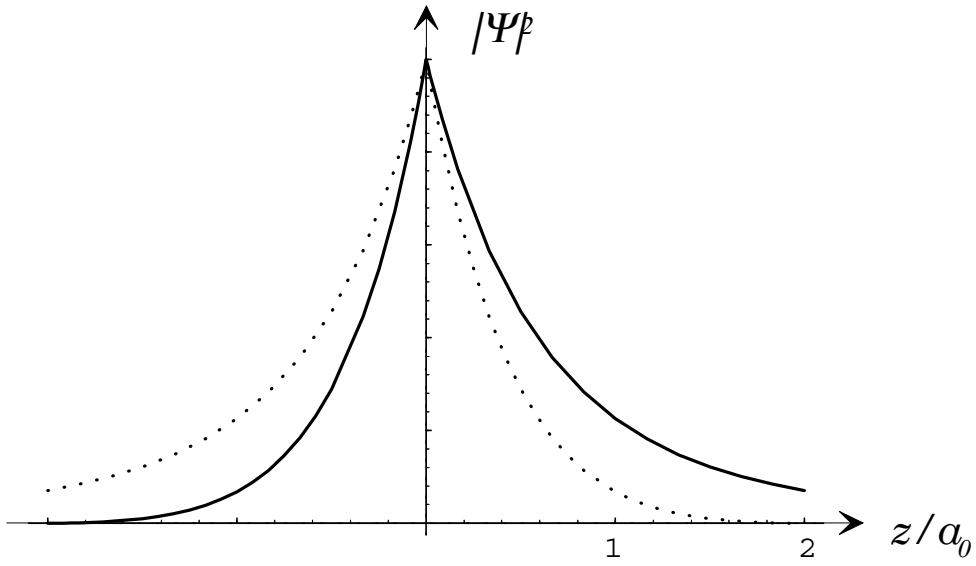

Figure 3.2: Carré du module de la fonction d'onde d'un atome d'hydrogène préparé dans une superposition quantique des états 1S et 2P, $m = 0$, à deux instants séparés par une demi–période de Bohr (en traits pleins et pointillés). On représente ce carré le long de l'axe de quantification Oz. On note que le barycentre de la distribution de charge oscille selon Oz.

moyenne oscille autour de l'origine à la fréquence de Bohr. Pour préparer une telle superposition d'états, nous ne pouvons guère, dans le cadre de ce modèle, que soumettre l'atome à une onde incidente, elle aussi traitée classiquement. Nous pourrons alors calculer le dipôle moyen en appliquant l'équation de Schrödinger et déterminer ensuite le champ rayonné. Essentiellement, nous nous cantonnerons au problème de la diffusion. La situation sera bien sûr très différente selon que l'onde incidente sera résonnante sur la fréquence de Bohr d'une transition autorisée ou non.

3.2.2 Diffusion du rayonnement

Nous soumettons donc notre atome quantique à une onde plane incidente, le champ électrique au voisinage de l'origine étant de la forme $\mathbf{E} = E_0\mathbf{u}_z \exp(i(kx - \omega t))$ (nous choisissons une onde plane polarisée selon $\mathbf{u}_z$ et se propageant dans la direction $\mathbf{u}_x$). Nous négligerons bien sûr la réaction de rayonnement, c'est à dire le champ rayonné par l'atome lui même, par rapport au champ incident. Il ne serait d'ailleurs pas possible d'en tenir compte de façon satisfaisante dans ce modèle. Le potentiel vecteur de cette onde s'écrit $\mathbf{A} = -(iE_0/\omega)\mathbf{u}_z \exp(i(kx - \omega t))$. Son potentiel scalaire est bien sûr identiquement nul. Nous nous placerons ici en jauge de Coulomb.

Hamiltonien d'interaction

La première étape du calcul est de déterminer la forme du Hamiltonien atomique en présence de l'onde incidente. Pour cela, nous nous appuierons sur les résultats de la partie sur la mécanique analytique, où nous avons établi l'expression du Hamiltonien classique d'une particule chargée en présence d'un champ. Nous remplacerons, en suivant les préceptes de la quantification canonique, les variables conjuguées $\mathbf{p}$ et $\mathbf{r}$ par des opérateurs conjugués $\mathbf{P}$ et $\mathbf{R}$ (nous noterons r la position au lieu de q).

Le Hamiltonien classique s'écrivant:

$$\mathcal{H} = \frac{(\mathbf{p} - q\mathbf{A}(\mathbf{r}))^2}{2m} + qV(\mathbf{r}) \ , \tag{3.41}$$

où $\mathbf{A}$ est le potentiel vecteur et V le potentiel scalaire, le Hamiltonien total quantique est simplement:

$$H = \frac{(\mathbf{P} - q\mathbf{A}(\mathbf{R}))^2}{2m} + qV(\mathbf{R}) \ , \tag{3.42}$$

obtenu en remplaçant position et impulsion par les opérateurs correspondants, y compris dans les expressions des potentiels. Dans le cas qui nous occupe, le potentiel vecteur est celui de l'onde plane incidente. Le potentiel scalaire est le potentiel électrostatique de noyau assurant la liaison de l'électron. Dans un calcul plus complet, pour un atome à plusieurs électrons, il conviendrait de tenir compte également de l'interaction entre électrons. Il faudrait aussi tenir compte d'autres termes dans le Hamiltonien, décrivant les structures fines ou hyperfines des niveaux.

En développant le carré et en n'oubliant pas que $\mathbf{P}$ et $\mathbf{A}$, devenu un opérateur comme fonction de $\mathbf{R}$, ne commutent pas, on obtient:

$$H = H_0 - \frac{q}{2m}(\mathbf{P}\cdot\mathbf{A} + \mathbf{A}\cdot\mathbf{P}) + \frac{q^2}{2m}\mathbf{A}^2 \ , \tag{3.43}$$

où $H_0 = \mathbf{P}^2/2m + V$ est le Hamiltonien de l'atome libre.

En général, les effets de l'onde incidente sont petits par rapport à l'effet du potentiel de liaison. En d'autres termes, le champ électrique de l'onde incidente est petit devant le champ électrostatique de liaison (de l'ordre de 10^{11} V/m). Les termes s'ajoutant à H_0 dans l'équation précédente ont donc une influence petite et les états propres de H sont essentiellement ceux de H_0. Nous pourrions donc traiter ce problème en utilisant le formalisme des perturbations de la mécanique quantique, perturbations dépendant du temps par l'intermédiaire de l'onde incidente. Plutôt que d'utiliser ce formalisme général, nous allons établir très simplement la forme de l'évolution de l'état atomique. Nous ferons pour cela, avant d'aller plus loin, deux approximations supplémentaires.

Si $\mathbf{A}$ est du premier ordre dans la perturbation que nous apportons à H_0, le dernier terme, proportionnel à $\mathbf{A}^2$, est du second ordre. Pour une intensité incidente suffisamment faible, il sera négligeable par rapport aux termes du premier ordre. En pratique, ce terme ne devient important que pour des intensités incidentes énormes, pour lesquelles le champ de l'onde est de l'ordre du champ de liaison. La seconde approximation est dite "approximation dipolaire". Tel que nous l'avons écrit, H est difficile à traiter en raison de la non–commutation de $\mathbf{P}$ et de $\mathbf{A}$. Notre problème est surtout intéressant pour traiter la diffusion d'une onde de fréquence voisine des transitions atomiques intenses, c'est à dire pour une onde dans le domaine visible. La longueur d'onde incidente, de l'ordre du micron, est alors beaucoup plus grande que la taille de l'atome. En première approximation, on peut assimiler le potentiel vecteur à sa valeur à l'origine, $\mathbf{A}(0)$. Cette approximation revient à ne tenir compte que du premier terme, dipolaire, dans le développement multipolaire de l'interaction de l'atome avec le champ. Avec cette approximation, le potentiel vecteur devient un simple nombre, indépendant de l'opérateur position. Il commute donc avec l'impulsion. On peut, avec ces deux approximations, très largement vérifiées en pratique, réécrire le Hamiltonien total sous la forme:

$$H = H_0 - \frac{q}{m}\mathbf{P}\cdot\mathbf{A}(0) \ . \tag{3.44}$$

En injectant finalement l'expression du potentiel vecteur de l'onde incidente, dont on prendra garde de ne garder que la partie réelle (il convient de n'utiliser que les champs physiques dans les Hamiltonien), on trouve:

$$H = H_0 + H_i \ , \tag{3.45}$$

avec

$$H_i = \frac{qE_0}{m\omega}P_z \sin(\omega t) \ . \tag{3.46}$$

Equations d'évolution des populations

Munis du Hamiltonien, nous allons maintenant calculer l'évolution temporelle de l'état atomique. De manière évidente, cet état peut s'écrire comme une superposition des états propres $|i\rangle$ du Hamiltonien non perturbé (ces états forment une base de l'espace de Hilbert). On écrira donc:

$$|\Psi(t)\rangle = \sum_i a_i(t)|i\rangle \ , \tag{3.47}$$

où la somme est à étendre à tous les états propres de H_0 (y compris les états du continuum) et où les a_f sont des amplitudes de probabilité complexes dépendant du temps. En portant directement ce développement dans l'équation de Schrödinger:

$$i\hbar\frac{\partial|\Psi\rangle}{\partial t} = H|\Psi\rangle \ , \tag{3.48}$$

on trouve:

$$i\hbar\sum_i \frac{da_i}{dt}|i\rangle = \sum_i a_i E_i|i\rangle + \frac{qE_0}{m\omega}\sum_i P_z \sin(\omega t) a_i|i\rangle \ . \tag{3.49}$$

Pour obtenir un système d'équations différentielles ordinaires, nous multiplierons scalairement les deux membres de cette équation par le "bra" $\langle j|$ où j est un niveau stationnaire arbitraire de l'atome non perturbé. En utilisant le fait que les états propres de H_0 sont orthogonaux, c'est à dire que

$$\langle j|i\rangle = \delta_{ji} \ , \tag{3.50}$$

on obtient, pour tout j,

$$\frac{da_j}{dt} = -i\omega_j a_j + \frac{qE_0}{im\omega\hbar}\sum_i \langle j|P_z|i\rangle a_i \sin(\omega t) \ , \tag{3.51}$$

où nous avons posé $\omega_j = E_j/\hbar$.

Nous pouvons simplifier un peu les écritures par un changement de variable. Le premier terme du second membre dans cette équation décrit l'évolution libre du coefficient a_j à la fréquence ω_j. On peut s'en affranchir en posant:

$$b_j = a_j e^{i\omega_j t} \ . \tag{3.52}$$

Ce changement de variable est en fait simplement un passage en représentation d'interaction par rapport au Hamiltonien non perturbé H_0. On obtient simplement alors l'équation d'évolution des b_j:

$$\frac{db_j}{dt} = \frac{qE_0}{im\omega\hbar}\sum_i \langle j|P_z|i\rangle \sin(\omega t) e^{i\omega_{ji}t} b_i(t) \ , \tag{3.53}$$

où $\omega_{ji} = \omega_j - \omega_i$ est la fréquence de Bohr de la transition de i vers j. Notons que cette quantité, algébrique, peut être négative.

Nous avons donc transformé l'équation de Schrödinger en un système différentiel ordinaire (les éléments de matrice $\langle j|P_z|i\rangle$ sont a priori facilement calculables). Il n'en reste pas moins que ce système comporte une infinité d'équations et qu'il n'est pas abordable dans le cas général. Nous pourrons aborder le problème dans deux situations limites seulement.

Diffusion non résonante

Nous considérerons ici la diffusion d'une onde non résonnante par un atome initialement dans son état fondamental g. Par non résonnante, nous désignons une onde dont la fréquence est très différente de toutes les fréquences de Bohr des transitions "autorisées" (telles que l'élément de matrice $\langle g|P_z|i\rangle$ soit non nul) partant de l'état g. Il conviendra aussi que son amplitude ne soit pas trop grande. Il est relativement naturel de considérer que cette perturbation non résonnante n'a qu'une très faible

probabilité d'exciter l'atome, c'est-à-dire de le transférer vers un état autre que g. Cela signifie que pratiquement toutes les amplitudes b_j restent petites (au moins du premier ordre dans l'amplitude de la perturbation, mesurée par E_0), sauf b_g qui est voisine de 1. Dans le second membre du système d'équations (3.51), le produit d'amplitudes du premier ordre par un terme lui aussi proportionnel à E_0 donne un terme négligeable du second ordre. Seul contribue à cette somme le niveau g.

L'évolution de b_j pour $j \neq g$ s'écrit alors:

$$\frac{db_j}{dt} = \frac{qE_0}{im\omega\hbar}\langle j|P_z|g\rangle \sin(\omega t)e^{i\omega_{jg}t} \ , \tag{3.54}$$

équation différentielle ordinaire qui s'intégre trivialement avec la condition initiale $b_j(0) = \delta_{jg}$. On trouve:

$$b_j(t) = \frac{qE_0}{2m\omega\hbar} i\langle j|P_z|g\rangle e^{i\omega_{jg}t}\left[\frac{e^{i\omega t} - e^{-i\omega_{jg}t}}{\omega + \omega_{jg}} - \frac{e^{-i\omega t} - e^{-i\omega_{jg}t}}{\omega_{jg} - \omega}\right] \ . \tag{3.55}$$

A partir de cette expression des b_j, nous pouvons facilement calculer le dipôle moyen $\langle \mathbf{D}\rangle = \langle \Psi(t)|\mathbf{D}|\Psi(t)\rangle$ avec $|\Psi\rangle = |g\rangle + \sum_j b_j(t)\exp(-i\omega_j t)|j\rangle$ (on remarquera que la norme de cet état n'est un qu'au premier ordre dans les b_j ce qui est conforme avec nos approximations). Après quelques manipulations algébriques sans intérêt, on trouve:

$$\langle \mathbf{D}\rangle = \mathbf{u}_z \frac{q^2E_0}{2m\omega\hbar}\left\{ i\sum_{j\neq g}\langle g|Z|j\rangle\langle j|P_z|g\rangle \left[\frac{e^{i\omega t} - e^{-i\omega_{jg}t}}{\omega + \omega_{jg}} - \frac{e^{-i\omega t} - e^{-i\omega_{jg}t}}{\omega_{jg} - \omega}\right] + c.c.\right\} \ . \tag{3.56}$$

Dans cette expression, apparaissent comme fréquences d'évolution la fréquence de l'onde incidente et la fréquence de Bohr de g vers j. Physiquement, nous pouvions nous attendre à ce que l'atome ne réponde qu'à la fréquence à laquelle il est forcé. En fait, la partie oscillant avec la fréquence de Bohr provient de deux imperfections du modèle. D'abord, nous avons supposé que le branchement de l'onde plane excitatrice intervenait instantanément à l'instant origine. Le spectre de la perturbation contient donc toutes les fréquences, en plus de la fréquence propre ω. Dans ce spectre, se trouvent en particulier les fréquences de Bohr, capables d'exciter de façon résonnante les transitions atomiques et de provoquer une réponse transitoire de l'atome. Comme nous avons également négligé tout amortissement, cette réponse transitoire s'étend indéfiniment dans le temps et contamine notre solution. On aurait tout à fait le même genre de comportement, comme on pourra s'en convaincre facilement, en branchant instantanément une perturbation non résonnante sur un oscillateur non amorti. Clairement, cette partie de la solution n'est pas physique. D'une part, la perturbation n'est jamais appliquée de façon soudaine (une prise en compte réaliste d'un branchement adiabatique de la perturbation réduirait considérablement le poids des fréquences de Bohr dans la solution). D'autre part, et surtout, nous n'avons introduit aucun amortissement dans notre problème. Il est clair que l'excitation des niveaux due au branchement soudain ne pourra pas durer plus longtemps que la durée de vie de ces niveaux si on prenait en compte l'émission spontanée. Nous ne pouvons tenir compte de façon convaincante de cet amortissement mais nous pouvons au moins supprimer dans l'expression du dipôle moyen les termes non physiques aux fréquences de Bohr. Nous aurons donc finalement:

$$\langle \mathbf{D}\rangle = \mathbf{u}_z \frac{q^2E_0}{2m\omega\hbar}\left\{ i\sum_{j\neq g}\langle g|Z|j\rangle\langle j|P_z|g\rangle \left[\frac{e^{i\omega t}}{\omega + \omega_{jg}} - \frac{e^{-i\omega t}}{\omega_{jg} - \omega}\right] + c.c.\right\} \ . \tag{3.57}$$

Notons ici un autre effet de l'absence d'amortissement. Les dénominateurs apparaissant dans cette expression présentent un comportement résonnant quand la fréquence excitatrice coïncide, au signe près, avec une des fréquences de Bohr. Dans ce cas, le dipôle moyen diverge, ce qui n'est pas physique. Il nous faut bien garder en mémoire que cette expression n'est valable que tant que les b_j restent petits, c'est à dire pour de petits dipôles moyens, loin de résonance et avec des ondes incidentes d'amplitude modérée.

Le produit d'éléments de matrice $\langle g|z|j\rangle\langle j|P_z|g\rangle$ est assez désagréable. Nous allons le transformer quelque peu. Nous pouvons en effet écrire, en utilisant les règles standard de commutation entre $\mathbf{R}$ et une fonction de $\mathbf{P}$:

$$[Z, H_0] = \left[Z, \frac{P_z^2}{2m}\right] = \frac{i\hbar}{m} P_z \tag{3.58}$$

(notons que V commute avec $\mathbf{R}$ et que les composantes x et y de $\mathbf{P}$ commutent avec Z). On en déduit immédiatement:

$$\langle j|P_z|g\rangle = \frac{m}{i\hbar}\langle j|ZH_0 - H_0Z|g\rangle = \frac{m}{i}\omega_{gj}\langle j|Z|g\rangle \tag{3.59}$$

En reportant cette expression dans celle du dipôle moyen, on obtient:

$$\langle \mathbf{D}\rangle = -\mathbf{u}_z \frac{q^2 E_0}{2\hbar} \sum_{j\neq g} \frac{\omega_{jg}}{\omega} |\langle j|Z|g\rangle|^2 \left[\frac{e^{i\omega t}}{\omega + \omega_{jg}} - \frac{e^{-i\omega t}}{\omega_{jg} - \omega}\right] + c.c. \tag{3.60}$$

Nous pouvons alors, pour faire le lien avec le modèle classique du paragraphe précédent, introduire une polarisabilité atomique "quantique" à la fréquence ω, $\alpha_q(\omega)$, définie par:

$$\mathbf{D} = \epsilon_0 \alpha_q E_0 \mathbf{u}_z \ . \tag{3.61}$$

En ne faisant apparaître que des parties réelles dans l'expression du dipôle, on a

$$\alpha_q(\omega) = \frac{2q^2}{\hbar\epsilon_0} \sum_j \frac{\omega_{jg}}{\omega_{jg}^2 - \omega^2} |\langle j|Z|g\rangle|^2 \tag{3.62}$$

(remarquons que la restriction $j \neq g$ dans la sommation a pu être éliminée, la moyenne de Z étant nulle dans le niveau g). Si nous rapprochons cette expression de la polarisabilité classique, déduite du modèle de Thomson de l'électron élastiquement lié de fréquence propre ω_0:

$$\alpha_c(\omega, \omega_0) = \frac{q^2}{m\epsilon_0} \frac{1}{\omega_0^2 - \omega^2} \ , \tag{3.63}$$

on voit que la polarisabilité quantique peut se mettre sous la forme d'une somme pondérée de polarisabilités classiques pour des oscillateurs chargés aux différentes fréquences de Bohr:

$$\alpha_q(\omega) = \sum_j f_{jg} \alpha_c(\omega, \omega_{jg}) \ , \tag{3.64}$$

où nous définissons la "force d'oscillateur" de la transition de g vers j, f_{jg} par:

$$f_{jg} = \frac{2m\omega_{jg}}{\hbar} |\langle j|Z|g\rangle|^2 \ . \tag{3.65}$$

Les forces d'oscillateur sont évidemment des quantités réelles, positives et sans dimension. On peut facilement aussi établir que la somme de toutes les forces d'oscillateur depuis le niveau g est égale à 1 (règle de Reich–Thomas–Kuhn). Pour cela, on remarque que:

$$f_{jg} = \frac{2m\omega_{jg}}{\hbar} \langle g|Z|j\rangle\langle j|Z|g\rangle \ . \tag{3.66}$$

En utilisant alors l'équation (3.59) établie plus haut, nous pouvons mettre la force d'oscillateur sous la forme:

$$f_{jg} = \frac{2}{i\hbar} \langle g|Z|j\rangle\langle j|P_z|g\rangle \ . \tag{3.67}$$

La relation de fermeture $\sum_j |j\rangle\langle j| = 1$ donne:

$$\sum_j f_{jg} = \frac{2}{i\hbar}\langle g|ZP_z|g\rangle = -\frac{2}{i\hbar}\langle g|P_z Z|g\rangle \ , \tag{3.68}$$

cette quantité, réelle, étant évidemment égale à son complexe conjugué. On a donc finalement:

$$\sum_j f_{jg} = \frac{1}{i\hbar}(\langle g|ZP_z - P_z Z|g\rangle) = 1 \ . \tag{3.69}$$

L'interprétation physique de l'expression de la polarisabilité quantique est donc transparente. Tout se passe comme si l'atome quantique était remplacé par une collection d'oscillateurs harmoniques, sans amortissement, chacun résonnant à une des fréquences de Bohr correspondant à toutes les transitions partant du niveau g. Chacun de ces oscillateurs contribue proportionnellement à sa force d'oscillateur. Celle-ci est nulle si l'élément de matrice de Z est nul. Les transitions de force d'oscillateur nulle sont dites "interdites"[7]. Dans beaucoup de cas, une des transitions originaires du niveau g porte l'essentiel de la force d'oscillateur. C'est par exemple le cas de la raie de résonance des alcalins, telle que la célèbre raie jaune du sodium, (qui est en fait un doublet en raison de la structure fine) qui rassemble presque toute la force d'oscillateur. Dans un tel cas, l'atome, pour ce qui est de la diffusion non résonnante, se comporte essentiellement comme un oscillateur chargé unique. Le modèle de Thomson est donc beaucoup plus réaliste que ne le laisse supposer sa simplicité. Ce que nous venons d'aborder dans le cadre d'une théorie atomique très simplifiée peut être généralisé à des systèmes quantiques plus complexes, comme des molécules. On peut tout à fait donner encore un sens à la notion de force d'oscillateur. C'est même le cas pour des systèmes macroscopiques, comme les systèmes excitoniques dans les semi–conducteurs.

Toutes les discussions du chapitre précédent sur les différents régimes de diffusion sont donc inchangées, à condition de tenir compte de toutes les transitions autorisées. On peut donc s'interroger sur la validité des approximations "haute fréquence" et "basse fréquence" en présence d'une infinité de raies de résonance. Le régime basse fréquence, de Rayleigh, correspond pour un atome à des fréquences incidentes plus petites que celles de la transition optique de plus basse énergie. Il peut exister des transitions d'énergie encore plus basse, entre niveaux hyperfins du fondamental, par exemple, mais elles ne correspondent pas à des transitions dipolaires électriques. La dépendance en ω_{jg} de la force d'oscillateur diminue également considérablement l'influence des transitions de très basse fréquence. Pour le régime haute fréquence, diffusion Thomson, le problème est a priori plus difficile. Il existe toujours une transition résonnante du fondamental vers le continuum, quelle que soit la fréquence incidente. On peut se rassurer en remarquant que les forces d'oscillateur associées à ces transitions sont extrêmement faibles. En effet, la fonction d'onde d'un électron libre d'énergie élevée est une onde quasi–plane de très courte longueur d'onde qui a un très mauvais recouvrement avec la fonction d'onde du fondamental, à symétrie sphérique et d'extension de l'ordre de a_0. On peut donc donner un sens au régime de diffusion de Thomson dès que la fréquence incidente est beaucoup plus grande que la fréquence nécessaire pour ioniser l'atome.

Entre les régimes de Thomson et de Rayleigh, on entre dans le domaine de la diffusion résonnante, que nous ne pouvons traiter avec ce modèle où on suppose constante la population du niveau g. La perturbation résonnante induira en effet des transitions entre les niveaux atomiques.

Interaction résonnante

Nous pourrons traiter, sans recourir à un modèle plus complet, le cas très particulier où l'onde incidente est strictement résonnante avec une transition atomique particulière, du niveau g vers le niveau j. On

[7] En fait, elles contribuent à la polarisabilité à un ordre plus élevé dans le développement multipolaire. Une transition interdite pour le rayonnement dipolaire électrique peut, par exemple, être autorisée pour le rayonnement quadripolaire. Comme nous l'avons vu dans les chapitres précédents, le rayonnement dipolaire est de loin le plus intense. La diffusion par un atome est donc en général largement dominée par les transitions dipolaires.

peut en effet supposer, comme l'onde incidente est non résonnante sur toutes les autres transitions (nous négligerons ici les problèmes liés à une éventuelle dégénérescence de deux niveaux) que seuls les niveaux j et g sont notablement peuplés, les autres étant essentiellement vides. Le système infini (3.53) est donc remplacé par un système de deux équations à deux inconnues, facilement résolu. On a en effet

$$\begin{aligned} \frac{db_j}{dt} &= \frac{qE_0}{i\hbar m\omega}\langle j|P_z|g\rangle \sin\omega t e^{i\omega_{jg}t} b_g \\ \frac{db_g}{dt} &= \frac{qE_0}{i\hbar m\omega}\langle g|P_z|j\rangle \sin\omega t e^{-i\omega_{jg}t} b_j \ . \end{aligned} \tag{3.70}$$

En utilisant le lien entre éléments de matrice de P_z et de Z (équation (3.59)), on peut faire apparaître dans les équations précédentes l'élément de matrice dipolaire:

$$d = q\langle j|Z|g\rangle = \frac{q}{im\omega}\langle j|P_z|g\rangle \ . \tag{3.71}$$

On peut remarque aussi qu'il intervient le produit de $\sin\omega t$ par $e^{i\omega_{jg}t}$. Pour $\omega_{jg} = \omega$ (c'est à dire à résonance), ce produit, en développant le sinus, fait intervenir la somme de deux termes, l'un constant et l'autre oscillant à 2ω. Si l'évolution des b est suffisamment lente, ce qui doit être le cas puisque nous sommes en représentation d'interaction par rapport aux énergies atomiques, on peut faire ce que les astronomes appellent une approximation séculaire. On néglige l'influence du terme rapide, en remarquant qu'il se moyenne à zéro en un temps court par rapport aux temps caractéristiques d'évolution des populations. En physique atomique, cette approximation porte le nom "d'approximation de l'onde tournante". En posant alors

$$\Omega = \frac{dE_0}{\hbar} \ , \tag{3.72}$$

on met le système différentiel sous la forme:

$$\frac{db_j}{dt} = \frac{\Omega}{2} b_g \ , \qquad \frac{db_g}{dt} = \frac{\Omega}{2} b_j \ , \tag{3.73}$$

dont la solution, assortie de la condition initiale $b_g = 1; b_j = 0$ est évidemment:

$$b_g = \cos\frac{\Omega}{2}t \qquad b_j = \sin\frac{\Omega}{2}t \ . \tag{3.74}$$

Les amplitudes de probabilité oscillent simplement en fonction du temps entre 0 et 1. Si nous calculons par exemple la probabilité de trouver l'atome dans l'état j, qui est proportionnelle à l'énergie atomique moyenne, nous trouverons:

$$P_j = \frac{1}{2}(1 - \cos\Omega t) \ . \tag{3.75}$$

Cette oscillation régulière entre les deux niveaux porte le nom "d'oscillation de Rabi", en l'honneur de Rabi qui la mit en évidence pour la première fois dans des expériences de résonance micro–onde sur des spins atomiques. La fréquence Ω est donc nommée "fréquence de Rabi".

A partir de cette expression des amplitudes de probabilité, il est aisé de calculer le dipôle électrique moyen. Il est évidemment égal à $2d\mathrm{Re}\,(\cos\frac{\Omega}{2}t \sin\frac{\Omega}{2}t \exp i\omega t)$. A des coefficients numériques près, le dipôle est donc proportionnel à $\sin\Omega t \cos\omega t$. Le dipôle moyen évolue donc à deux fréquences:

$$\omega \pm \Omega \ . \tag{3.76}$$

Notons que, pour les amplitudes incidentes où ce modèle est réaliste, on a toujours $\omega \gg \Omega$. Le spectre de la lumière émise par l'atome consiste donc en deux raies, symétriquement disposées par rapport à la fréquence atomique "nue". On peut comprendre aisément ce spectre en termes de "bandes latérales".

Le dipôle atomique, oscillant naturellement à la fréquence ω, est modulé en amplitude à la fréquence Ω. Dans le modèle simple que nous traitons ici, cette modulation est complète. On sait bien que le spectre est alors composé de deux bandes latérales à $\omega \pm \Omega$ (dans le cas d'une modulation partielle, il demeure une composante à la fréquence de la "porteuse" ω). En fait, nous savons que notre modèle est incomplet. Il ne tient pas compte de l'émission spontanée. Quand on l'inclut correctement, on constate que le spectre contient une troisième composante à la fréquence atomique ω dont l'amplitude est double de celles des "bandes latérales" précédentes. Le spectre du rayonnement d'un atome unique doit donc être constitué de trois raies, le célèbre "triplet de Mollow". Ce triplet peut facilement être mis en évidence en irradiant un échantillon atomique par un laser. Il faut en effet que la fréquence de Rabi, Ω, soit grande par rapport à la largeur des raies spectrales, déterminées par la durée de vie radiative du niveau excité ou des effets parasites comme l'effet Doppler dû à l'agitation thermique. Cela impose l'utilisation de lasers largement saturants, au sens où nous avons défini la saturation pour la diffusion résonnante dans le modèle classique. De nombreux effets intéressants d'optique quantique sont liés à ce régime d'oscillation de Rabi et "d'habillage" de l'atome par le champ d'un laser intense.

Notre modèle serait aussi insuffisant pour décrire l'oscillation de Rabi dans un très petit champ. Si celui ci ne contient que quelques photons, la quantification de l'énergie et donc de l'amplitude doit entrer en ligne de compte. Au lieu d'une fréquence de Rabi unique, variant continûment avec l'amplitude classique du champ électromagnétique, on peut s'attendre à ce que l'oscillation s'effectue à des fréquences discrètes, correspondant aux différents nombres de photons présents dans le champ.

Dans la plupart des situations ordinaires de couplage d'un atome à un champ laser, cette quantification de la fréquence de Rabi passe complètement inaperçue. Pour que l'oscillation s'effectue plus rapidement que n'agissent les différentes causes de relaxation, il faut en effet que le champ contienne un nombre énorme de photons, très rapidement renouvelés. La différence entre deux nombres de photons consécutifs est si petite par rapport à ces nombres que l'oscillation de Rabi apparaît encore comme une sinusoïde, avec une fréquence proportionnelle à l'amplitude moyenne du champ électromagnétique, variant continûment. Pour observer la quantification, il faut se placer dans des conditions très particulières où un atome est fortement couplé au champ de quelques photons seulement (en d'autres termes, telles que l'intensité de saturation de la transition atomique soit de l'ordre du photon). Il faut de plus que le nombre de photons reste bien constant pendant toute la durée de l'oscillation.

On peut réaliser ce genre de situation dans le contexte de ce qu'il est maintenant convenu d'appeler "l'électrodynamique quantique en cavité". Un atome, très fortement couplé au rayonnement, est placé dans une cavité résonnante de haute surtension. L'atome est par exemple un "état de Rydberg circulaire" (voir l'annexe de la partie I sur le modèle de Bohr). Il se comporte comme une antenne géante pour le rayonnement millimétrique. De plus, la durée de vie de ces niveaux circulaires, en dépit de leur caractère exotique, est très longue. Enfin, on peut détecter ces atomes de manière sélective et sensible. L'atome, préparé dans un jet atomique, traverse une cavité résonnante sur une transition vers le niveau circulaire immédiatement inférieur. La transition étant dans le domaine des longueurs d'onde millimétriques, on peut avoir une cavité d'excellente qualité en utilisant des miroirs supraconducteurs. La cavité joue alors le rôle d'une "boîte à photons" conservant le champ pendant un temps beaucoup plus long que le temps de transit de l'atome à travers la cavité. Dans ces conditions, on peut réaliser une situation de "couplage fort", où toute dissipation est négligeable par rapport au couplage de l'atome avec la cavité et où le champ d'un seul photon est suffisant pour saturer la transition atomique. Notons que l'expérience doit s'effectuer à très basse température (moins de 1K) pour que le rayonnement du corps noir, particulièrement important dans le domaine millimétrique, ne remplisse pas la cavité.

L'énergie d'un champ de n photons à la fréquence ω étant $n\hbar\omega$, les amplitudes autorisées du champ dans la cavité et donc les fréquences de Rabi possibles, doivent être proportionnelles aux racines carrées successives des entiers. Un calcul complètement quantique du couplage d'un atome à deux niveaux avec un seul mode du champ électromagnétique confirme tout à fait cette approche intuitive. Ce qui est moins intuitif a priori, c'est que l'expression de la fréquence de Rabi dans le champ de n

photons est $\Omega\sqrt{n+1}$. Elle est non nulle même si la cavité est vide. En fait, l'oscillation qu'on observe quand l'atome entre dans une cavité vide est ce qui reste, dans ces conditions très exceptionnelles, de l'émission spontanée. L'atome excité émet un photon dans la cavité. Si cette émission se produisait dans l'espace libre, le photon s'échapperait à la vitesse de la lumière et l'atome resterait dans le niveau inférieur de la transition. Dans la cavité, en revanche, le photon reste piégé au voisinage de l'atome, prêt à être réabsorbé, puis réémis... Dans une telle cavité, l'émission spontanée devient un phénomène réversible, oscillatoire!

Dans un champ quantique ordinaire, tel que celui produit par une source classique (mettant en jeu des courants oscillants macroscopiques) faiblement couplée à la cavité , le nombre de photons n'est pas parfaitement défini. Il obéit en fait à une statistique de Poisson, avec une variance de l'ordre de la valeur moyenne. On montre alors que l'oscillation de Rabi s'effectue simultanément à toutes les fréquences correspondant aux nombres de photons présents dans cette distribution statistique, chaque fréquence apparaissant dans la transformée de Fourier du signal avec un poids égal à la probabilité $p(n)$ du nombre de photons.

Les résultats d'une expérience récente[8] réalisée selon ce principe sont présentés sur la figure 3.3. La première colonne présente les signaux d'oscillations de Rabi observés en fonction du temps. La courbe du haut correspond à une cavité vide. On y observe l'émission spontanée oscillante. Les courbes suivantes correspondent à des champs d'amplitude moyenne croissante dans la cavité. On observe bien que le signal n'est plus du tout sinusoïdal. Il présente une structure complexe, avec des phases de latence et de renaissance des oscillations, qu'aucun modèle semi–classique ne permet de comprendre. La deuxième colonne présente les transformées de Fourier de ces signaux. Pour la cavité vide, on observe une fréquence pure, $\Omega/2\pi = 47$ kHz, correspondant bien aux prédictions théoriques pour l'émission spontanée oscillante. Quand un champ est injecté dans la cavité, on observe des pics discrets dans cette transformée de Fourier dont les positions s'échelonnent bien comme les racines carrées des entiers successifs.

L'atome, dans cette expérience, est en quelque sorte une sonde qui mesure l'amplitude du champ électromagnétique. Les fréquences discrètes observées dans le spectre du signal de Rabi sont donc une preuve très directe de la quantification de l'amplitude et donc de l'énergie, du champ dans une cavité. Il y eut, depuis le début du siècle, de nombreuses preuves de la quantification du champ électromagnétique. L'émission spontanée, l'effet Compton, le déplacement de Lamb sont les plus anciennes. Tout le développement de l'optique quantique moderne, où l'on peut manipuler les fluctuations quantiques des champs, repose sur cette notion et la confirme avec éclat. Cependant, la manifestation la plus simple de la quantification du rayonnement, l'aspect discret de l'énergie dans une "boîte à photons" avait toujours échappé à l'expérience[9]. Il fallait en effet disposer d'un détecteur très particulier pour révéler ce comportement.

Notons enfin, dans la figure 3.3 la troisième colonne, qui présente les poids des différentes composantes de fréquences et donc les probabilités des nombres de photons correspondant. Elles s'accordent très bien avec la loi de Poisson attendue. Cela permet de déterminer avec précision le nombre moyen de photons dans la cavité.

---

[8] M. Brune, F. Schmidt-Kaler, A. Maali, J. Dreyer, E. Hagley, J.M. Raimond, S. Haroche, Phys. Rev. Lett. **76**,1800 (1996): "Quantum Rabi oscillation: a direct test of field quantization in a cavity". On pourra aussi consulter, sur les principes généraux de l'électrodynamique en cavité S. Haroche, J.M. Raimond, Scientific American, Avril 1993: "Cavity Quantum Electrodynamics"; Traduction française dans "Pour la Science", Juin 1993: "Electrodynamique Quantique en Cavité"; Réimprimé dans le numéro spécial "La physique Quantique", Pour la Science, Juin 1994. On pourra enfin consulter, pour des expériences plus récentes S. Haroche, J.M. Raimond et M. Brune, la Recherche, Sept. 1997.

[9] Les "clics" discrets d'un détecteur fondé sur l'effet photoélectrique ne sont pas une preuve absolue de la quantification du rayonnement. Dès 1926, à l'aube de la mécanique quantique moderne, Dirac et Wentzel ont montré que toutes les caractéristiques de l'effet photoélectrique pouvaient s'obtenir en couplant un champ classique à un détecteur quantifié (un atome par exemple). Einstein lui même considérait l'effet photoélectrique comme une conséquence importante de la notion de photon mais il avait fondé sa preuve de la quantification de l'énergie du champ sur des notions beaucoup plus fondamentales.

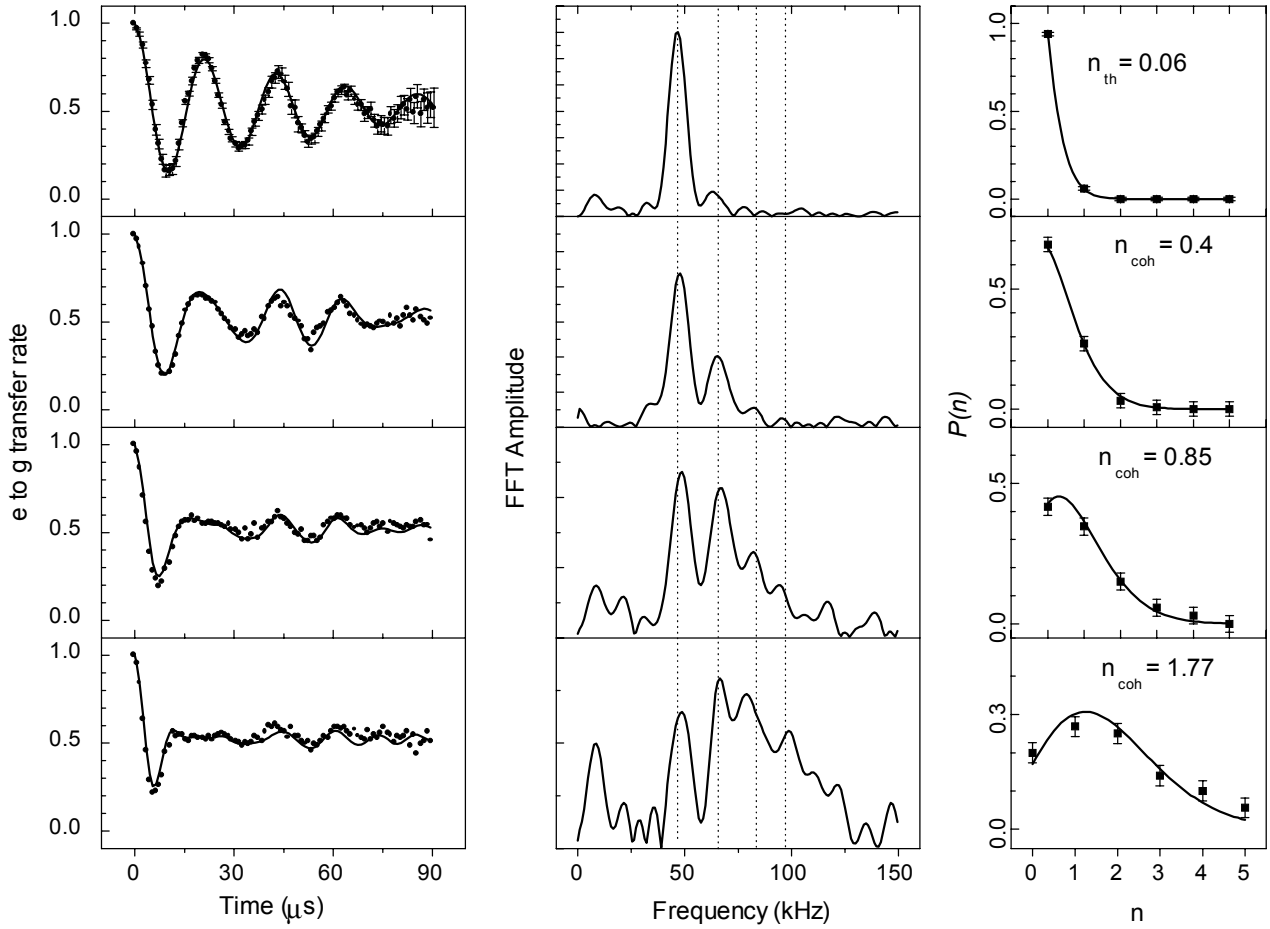

Figure 3.3: La première colonne présente les signaux d'oscillation de Rabi quantique obtenus en couplant un atome de Rydberg circulaire à une cavité millimétrique supraconductrice contenant un champ de quelques photons (l'amplitude moyenne de ce champ est croissante de haut en bas, la cavité est vide pour la courbe du haut). La deuxième colonne présente les transformées de Fourier de ces signaux, faisant clairement apparaître la quantification de l'amplitude du champ. Enfin, la troisième colonne présente les poids des composantes de fréquence ou des nombres de photons, obéissant à une loi de Poisson. Les nombres moyens de photons (0.40 ($\pm$0.02), 0.85 ($\pm$0.04) et 1.77 ($\pm$0.15) pour les trois lignes du bas) s'en déduisent. Pour la ligne supérieure, on observe un petit nombre de photons moyen (0.06) qui correspond au champ thermique résiduel dans la cavité à la température de l'expérience.

3.3 Diffusion par un milieu dense

Nous nous attacherons dans ce paragraphe au calcul du champ et de l'intensité diffusés par une assemblée d'atomes, ou plus généralement de centres diffuseurs. Chacun de ces centres sera, dans ce problème, entièrement défini par sa polarisabilité. Nos calculs pourront donc s'appliquer à une très grande variété de centres diffuseurs et, en particulier, aux atomes classiques ou semi–classiques traités dans les paragraphes précédents.

Nous essaierons en particulier de comprendre un apparent paradoxe. Nous avons vu en effet, dans le paragraphe sur la diffusion Rayleigh, que les atomes devaient efficacement diffuser un rayonnement incident, d'autant plus efficacement que la fréquence est élevée. Nous avons expliqué ainsi la diffusion du rayonnement bleu par le ciel. On peut cependant a priori s'étonner de ce que la diffusion par ce milieu, relativement peu dense, soit beaucoup plus efficace que la diffusion par un milieu transparent solide, contenant un nombre beaucoup plus grand de centres diffuseurs. Nous avons en effet montré que la longueur d'extinction typique (la longueur sur laquelle le rayonnement incident est notablement affaibli) est de quelques dizaines de kilomètres pour l'atmosphère. Pour des fibres optiques de haute qualité, la longueur d'atténuation est sensiblement plus grande, alors que la densité du milieu est au moins mille fois plus élevée. Il semble donc qu'un milieu dense à l'échelle de λ ne diffuse de manière particulièrement peu efficace. Nous essaierons aussi de comprendre pourquoi les cristaux, éclairés par une longueur d'onde de l'ordre de la maille cristalline, ne diffusent que dans des directions privilégiées et comment ces directions sont reliées aux paramètres du réseau cristallin.

Pour obtenir ces résultats nous ferons dans ce chapitre un certain nombre d'approximations. Elles seront détaillées dans le prochain paragraphe. Notons dès maintenant que nous ne traiterons que de la diffusion par un milieu statique. Un milieu où les centres diffuseurs se déplacent, même éclairé par une onde monochromatique, ne diffuse pas uniquement à la fréquence incidente en raison de l'effet Doppler. Pour traiter ce cas en détails, il nous faudrait des outils d'analyse d'un champ non monochromatique (fonctions de corrélation en particulier) qui ne seront pas introduites dans ce cours. Nous indiquerons brièvement à la fin du chapitre comment nos résultats seraient qualitativement modifiés dans ce cas. Nous donnerons une application importante à la diffusion par un milieu dépendant du temps avec la diffusion Brillouin (diffusion d'une onde électromagnétique par une onde sonore).

3.3.1 Notations. Champ diffusé.

Nous considérons donc un milieu comprenant des centres diffuseurs, que nous appellerons "atomes", localisés au voisinage de l'origine (voir figure 3.4). Les diffuseurs, en nombre total $\mathcal{N}$, indicés par l'indice i, sont situés en $\mathbf{r}_i$. La polarisabilité de chaque diffuseur sera notée $\alpha_i(\omega)$. Elle peut en effet dépendre de la fréquence du rayonnement incident. Dans presque tous les calculs, par la suite, on n'explicitera pas sa dépendance en ω. On observe le rayonnement dans la direction $\mathbf{n}$, définie par les angles d'Euler θ et ϕ. L'observateur est à une distance R, très grande par rapport à la taille du milieu et par rapport à toutes les autres dimensions caractéristiques du problème.

Le milieu est éclairé par une onde plane monochromatique, de fréquence ω, se propageant selon Oz, polarisée selon Ox et d'amplitude E_0. On pourrait traiter le cas d'une polarisation incidente arbitraire en utilisant le principe de superposition. On pourrait traiter également le cas d'une onde de comportement temporel quelconque en la décomposant en composantes de Fourier. On posera $\mathbf{k}_0 = k\mathbf{u}_z$, avec $k = \omega/c$ et $\mathbf{E}_0 = E_0\mathbf{u}_x$.On supposera que tous les centres diffuseurs sont soumis à l'onde incidente non modifiée. On suppose donc que la diffusion n'atténue pas l'onde incidente, c'est à dire que le milieu diffuseur n'est pas trop dense (c'est bien sûr une densité de polarisabilité qui importe ici). Nous supposerons également que la diffusion multiple est négligeable. Nous ne tiendrons donc pas compte, dans le calcul des dipôles induits, des champs rayonnés par les autres dipôles. Là encore, il s'agit d'une approximation de faible densité.

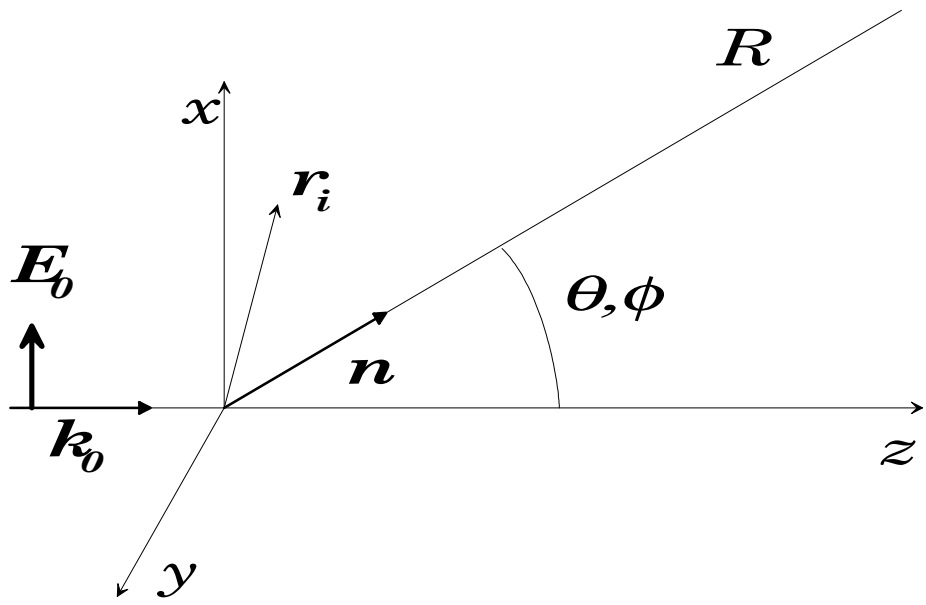

Figure 3.4: Notations pour la diffusion par un milieu matériel

L'onde incidente "vue" par l'atome i peut donc s'écrire:

$$\mathbf{E}_0 e^{i(\mathbf{k}_0\cdot\mathbf{r}_i-\omega t)} \ . \tag{3.77}$$

L'atome i prend donc un dipôle électrique égal à:

$$\mathbf{d}_i = \epsilon_0\alpha_i\mathbf{E}_0 e^{i(\mathbf{k}_0\cdot\mathbf{r}_i-\omega t)} \ . \tag{3.78}$$

En utilisant les résultats sur le champ du dipôle, on peut écrire le champ rayonné par l'atome i au niveau de l'observateur (le vecteur joignant l'atome à l'observateur est évidemment $R\mathbf{n}-\mathbf{r}_i$) comme:

$$\mathbf{E}_i = \frac{1}{4\pi}\frac{e^{ik|R\mathbf{n}-\mathbf{r}_i|}}{|R\mathbf{n}-\mathbf{r}_i|}\frac{\omega^2}{c^2}\alpha_i e^{i(\mathbf{k}_0\cdot\mathbf{r}_i-\omega t)}(\mathbf{n}\times\mathbf{E}_0)\times\mathbf{n} \ . \tag{3.79}$$

Pour aller plus loin, nous allons utiliser le fait que la distance d'observation est très grande: $R \gg |\mathbf{r}_i|$. Nous allons donc, comme nous l'avons déjà fait souvent, traiter, dans le terme d'onde sphérique, le dénominateur d'amplitude à l'ordre zéro en $R/|\mathbf{r}_i|$ et le terme de phase à l'ordre 1. On a:

$$|R\mathbf{n}-\mathbf{r}_i| \simeq R - \mathbf{n}\cdot\mathbf{r}_i \ , \tag{3.80}$$

et donc:

$$\mathbf{E}_i = \frac{1}{4\pi}\frac{e^{ikR}}{R}\frac{\omega^2}{c^2}\alpha_i e^{i((\mathbf{k}_0-\mathbf{k})\cdot\mathbf{r}_i)}e^{-i\omega t}(\mathbf{n}\times\mathbf{E}_0)\times\mathbf{n} \ , \tag{3.81}$$

où l'on a posé:

$$\mathbf{k} = k\mathbf{n}. \tag{3.82}$$

On peut finalement mettre ce résultat sous la forme:

$$\mathbf{E}_i = \mathbf{e}_d\alpha_i e^{-i\mathbf{q}\cdot\mathbf{r}_i} \ , \tag{3.83}$$

avec

$$\mathbf{q} = \mathbf{k} - \mathbf{k}_0 \ . \tag{3.84}$$

$\mathbf{q}$ apparaît donc comme la variation de vecteur d'onde entre l'onde incidente et l'onde diffusée. Il définit bien sûr complètement la direction d'observation. A un facteur $\hbar$ près, $\mathbf{q}$ est aussi la différence de quantité de mouvement entre les photons incidents et les photons diffusés. Notons que dans beaucoup d'autres cas (collisions de particules..), les amplitudes de diffusion s'expriment simplement à partir de ce transfert d'impulsion. On a aussi posé

$$\mathbf{e}_d = \frac{1}{4\pi}\frac{e^{ikR}}{R}\frac{\omega^2}{c^2}e^{-i\omega t}(\mathbf{n}\times\mathbf{E}_0)\times\mathbf{n} \ . \tag{3.85}$$

Notons à ce point que $\mathbf{e}_d$ est indépendant du diffuseur et de sa position. Il s'agit donc d'une donnée simple de la géométrie de la diffusion. Elle contient toute la dépendance angulaire du diagramme de rayonnement du dipôle. Elle s'annule en particulier pour une direction d'observation selon Ox. Tous les dipôles sont polarisés selon Ox et ne sauraient donc rayonner dans cette direction. $\mathbf{e}_d$ peut s'exprimer simplement en fonction des angles d'Euler de la direction d'observation:

$$\mathbf{e}_d = \frac{1}{4\pi}\frac{e^{ikR}}{R}\frac{\omega^2}{c^2}e^{-i\omega t}E_0(1-\sin^2\theta\cos^2\phi) \ . \tag{3.86}$$

Pour comparer ce résultat avec celui que nous avions établi dans le paragraphe sur le rayonnement du dipôle, on prendra garde que le dipôle est ici aligné avec l'axe $0x$, alors que les angles d'Euler sont relatifs à l'axe Oz, direction de propagation de l'onde incidente.

Le champ total rayonné par le milieu résulte de l'addition cohérente des champs diffusés par tous les centres

$$\mathbf{E} = \sum_i \mathbf{E}_i \ . \tag{3.87}$$

On peut passer facilement de cette somme discrète à une intégrale continue sur le volume V du milieu en introduisant une densité de polarisation. En fait, nous poserons:

$$\delta n(\mathbf{r}) = \frac{1}{2}\sum_i \alpha_i\delta(\mathbf{r}-\mathbf{r}_i) \ . \tag{3.88}$$

Nous verrons en effet dans la prochaine partie que l'indice de réfraction d'un milieu suffisamment peu dense (au sens de la polarisation) est l'unité plus le demi produit de la polarisabilité atomique par la densité numérique du milieu (le résultat est évidemment sans dimension). Comme notre milieu est supposé faiblement diffusant, δn apparaît clairement relié à l'écart à un de l'indice de réfraction. Notons toutefois que l'indice habituel n'a de sens que si on moyenne les densités de polarisation atomique à une échelle grande par rapport à la distance moyenne entre atomes et petite devant λ. Nous n'avons pas réalisé ce moyennage ici, et δn est une quantité tenant compte de la position individuelle de tous les atomes.

Avec ces notations, nous obtenons le résultat central de ce chapitre: le champ total diffusé dans la direction définie par $\mathbf{q}$, se met sous la forme:

$$\mathbf{E}(\mathbf{q}) = 2\mathbf{e}_d\int \delta n(\mathbf{r})e^{-i\mathbf{q}\cdot\mathbf{r}}\,d\mathbf{r} \ . \tag{3.89}$$

Nous n'indiquons pas les bornes de l'intégrale, δn étant nul en dehors du volume V du milieu. Les propriétés de la diffusion sont donc décrites par la transformée de Fourier de la distribution d'indice du milieu. Ce résultat n'est pas sans évoquer la diffraction de Fraunhofer. La différence essentielle est que nous traitons ici de sources réelles réparties en volume et non des sources fictives du principe de Huygens, réparties sur la surface d'une ouverture.

Dans la suite, nous utiliserons surtout cette expression du champ diffusé. On peut noter toutefois qu'il est possible d'exprimer simplement l'intensité totale diffusée dans la direction $\mathbf{q}$. La structure de l'onde à grande distance étant très voisine de celle d'une onde plane, on peut écrire en effet:

$$I = \frac{\epsilon_0 c}{2}\mathbf{E}\cdot\mathbf{E}^* = 2\epsilon_0 c|\mathbf{e}_d|^2\int \delta n(\mathbf{r})e^{-i\mathbf{q}\cdot\mathbf{r}}\,d\mathbf{r}\int \delta n^*(\mathbf{r}')e^{i\mathbf{q}\cdot\mathbf{r}'}\,d\mathbf{r}' \ . \tag{3.90}$$

En posant $\mathbf{r}' = \mathbf{r} - \boldsymbol{\rho}$, le produit de transformées de Fourier peut se mettre sous la forme d'un produit de convolution:

$$I = 2\epsilon_0 c|\mathbf{e}_d|^2 \int \delta n(\mathbf{r})\delta n^*(\mathbf{r}-\boldsymbol{\rho})e^{i\mathbf{q}\cdot\boldsymbol{\rho}}\, d\mathbf{r}d\boldsymbol{\rho} \ . \tag{3.91}$$

On peut donc mettre finalement l'intensité diffusée sous la forme d'une transformée de Fourier inverse en $\boldsymbol{\rho}$:

$$I = 2\epsilon_0 c|\mathbf{e}_d|^2 \int \Gamma(\boldsymbol{\rho})e^{i\mathbf{q}\cdot\boldsymbol{\rho}}\, d\boldsymbol{\rho} \ , \tag{3.92}$$

où la fonction $\Gamma(\boldsymbol{\rho})$, définie par:

$$\Gamma(\boldsymbol{\rho}) = \int \delta n(\mathbf{r})\delta n^*(\mathbf{r}-\boldsymbol{\rho})d\mathbf{r} \tag{3.93}$$

est la fonction d'autocorrélation spatiale de la répartition d'indice du milieu. L'intensité diffusée est donc la transformée de Fourier spatiale de la fonction d'autocorrélation de l'indice. Ce résultat se généralise aux milieux diffuseurs dépendant du temps. C'est alors la densité spectrale d'énergie diffusée qui est égale à la transformée de Fourier spatio–temporelle de la fonction de corrélation spatio–temporelle de l'indice.

3.3.2 Cas d'un milieu homogène

Considérons d'abord la diffusion par un milieu dense homogène, comme par exemple un verre transparent. Les distances moyennes entre centres diffuseurs, de l'ordre de l'Ångström, sont alors beaucoup plus petites que la longueur d'onde. On peut donc remplacer, dans la transformée de Fourier donnant le champ diffusé, le terme $\delta n(\mathbf{r})$ par une constante à l'intérieur du volume V – nous justifierons rigoureusement ce remplacement dans le prochain paragraphe. Nous poserons $n = 1 + \delta n$. Pour un milieu dense et homogène, n est effectivement l'indice de réfraction ordinaire [10].

Le champ diffusé s'écrit alors:

$$\mathbf{E} = 2\mathbf{e}_d(n-1)\int_V e^{-i\mathbf{q}\cdot\mathbf{r}}\, d\mathbf{r} \tag{3.94}$$

Dans la limite d'un volume infini, l'intégrale serait, à un facteur près, une fonction de Dirac $\delta(\mathbf{q})$. L'intégrale, ne portant que sur un volume fini, reste finie. Elle n'est toutefois non négligeable que dans un domaine étroit autour de $\mathbf{q} = 0$ (dont l'extension en termes de vecteurs d'onde est de l'ordre de l'inverse des dimensions transverses du milieu – un résultat élémentaire de théorie de la diffraction). Pour $\mathbf{q} = 0$, la valeur de l'intégrale est V. On peut donc écrire:

$$\int_V e^{-i\mathbf{q}\cdot\mathbf{r}}\, d\mathbf{r} = V\delta_V(\mathbf{q}) \ , \tag{3.95}$$

où δ_V est une fonction très piquée au voisinage de l'origine, de valeur maximale 1.

Le milieu homogène diffuse donc un champ proportionnel au volume, ou encore au nombre d'atomes $\mathcal{N}$, seulement dans la direction initiale de l'onde incidente. Il n'y a donc pas, dans les milieux denses et homogènes, de diffusion dans des directions latérales. On comprend ainsi la très faible atténuation par diffusion dans les fibres optiques. La diffusion vers l'avant, cohérente avec l'onde incidente, interfère avec celle-ci. C'est de cette interférence que résulte l'onde se propageant finalement dans le milieu. Le calcul complet des caractéristiques de cette onde à partir de l'expression du champ diffusé est complexe. Nous verrons dans la prochaine partie comment établir rigoureusement et simplement ce résultat en utilisant la théorie de la réponse linéaire. Nous montrerons que l'onde totale est une onde plane se propageant avec un vecteur d'onde $n\mathbf{k}_0$ et une vitesse de phase c/n. L'indice de réfraction du milieu peut donc être compris comme résultant de la diffusion cohérente vers l'avant de tous les atomes du milieu.

[10] La valeur constante de δn étant bien sûr invariante dans les procédures de moyennage qui nous seront nécessaires pour définir l'indice de réfraction.

3.3.3 Diffusion par un cristal

Considérons maintenant plus complètement le problème de la diffusion par une assemblée dense d'atomes immobiles. Nous allons supposer que ces atomes sont aux nœuds d'un réseau cristallin. Pour simplifier les calculs, nous prendrons un réseau cubique simple, les directions principales du cristal correspondant aux axes. La généralisation à d'autres orientations cristallines ou à d'autres types de réseaux cristallins ou quasi–cristallins ne pose que des problèmes algébriques. La position des atomes, supposés tous identiques, est donc définie par trois entiers, n_x, n_y et n_z, avec:

$$\mathbf{r}_i = a(n_x\mathbf{u}_x + n_y\mathbf{u}_y + n_z\mathbf{u}_z) \ , \tag{3.96}$$

où a est la maille cristalline. Nous noterons N_x, N_y et N_z les nombres d'atomes dans les trois directions de l'espace (nous supposons donc que la forme globale du cristal est aussi un cube). On a évidemment $\mathcal{N} = N_xN_yN_z$. Pour ce problème, il est bien sûr avantageux d'utiliser l'expression du champ diffusé comme une somme discrète sur la position des atomes individuels:

$$\mathbf{E} = \mathbf{e}_d\alpha \sum_{n_x,n_y,n_z} e^{-ia\mathbf{q}\cdot(n_x\mathbf{u}_x+n_y\mathbf{u}_y+n_z\mathbf{u}_z)} \ . \tag{3.97}$$

Ce champ peut encore s'écrire en termes des composantes q_x, q_y et q_z du transfert de vecteur d'onde:

$$\mathbf{E} = \mathbf{e}_d\alpha \left(\sum_{n_x=1}^{N_x} e^{iq_xn_xa}\right)\left(\sum_{n_y=1}^{N_y} e^{iq_yn_ya}\right)\left(\sum_{n_z=1}^{N_z} e^{iq_zn_za}\right) \ . \tag{3.98}$$

Les sommes de séries géométriques s'évaluent sans difficulté. A un facteur de phase sans intérêt près,

$$\sum_{n_x=1}^{N_x} e^{iq_xn_xa} = \frac{\sin\frac{q_xN_xa}{2}}{\sin\frac{q_xa}{2}} \ . \tag{3.99}$$

Le champ diffusé est donc le produit de trois fonctions de diffraction simples portant sur les composantes de $\mathbf{q}$. Les directions où l'intensité diffusée est maximale sont telles que, simultanément:

$$q_x = \frac{2\pi}{a}p_x, \qquad q_y = \frac{2\pi}{a}p_y, \qquad q_z = \frac{2\pi}{a}p_z \ , \tag{3.100}$$

où p_x, p_y, p_z sont trois entiers. Le champ diffusé n'est donc maximal que dans des directions discrètes pour lesquelles le vecteur d'onde $\mathbf{q}$ est sur un point d'un réseau cubique de maille $2\pi/a$. Ce réseau, qui est essentiellement le réseau obtenu par transformée de Fourier spatiale du réseau initial, est appelé réseau réciproque. Pour tout réseau cristallin, il existe un réseau réciproque. Ce n'est toutefois que dans le cas du réseau cubique simple que le réseau réciproque est géométriquement identique au réseau initial. Notons à ce point que l'amplitude du champ diffusé dans un de ces maximums est proportionnelle à $N_xN_yN_z = \mathcal{N}$. Il s'agit donc bien, comme dans le cas du milieu homogène, d'une diffusion cohérente où le champ est proportionnel au nombre d'atomes et l'intensité au carré de ce nombre.

La géométrie de la diffusion est représentée sur la figure 3.5. Le vecteur $\mathbf{q}$ appartient au réseau réciproque. Les vecteurs d'onde incident et émergent, $\mathbf{k}_0$ et $\mathbf{k}$ sont donc tels que leur différence est un point du réseau réciproque. Une solution est $\mathbf{q} = 0$ correspondant à la diffusion vers l'avant. Elle est bien sûr toujours possible (quelle que soit la nature du réseau cristallin, évidemment). Les normes des vecteurs $\mathbf{k}_0$ et $\mathbf{k}$ étant identiques (ω/c), il n'existe pas nécessairement d'autre solution. Notons en effet Θ l'angle entre $\mathbf{k}_0$ et $\mathbf{k}$ (angle de diffusion). On a évidemment en module:

$$q = 2k\sin\frac{\Theta}{2} \ . \tag{3.101}$$

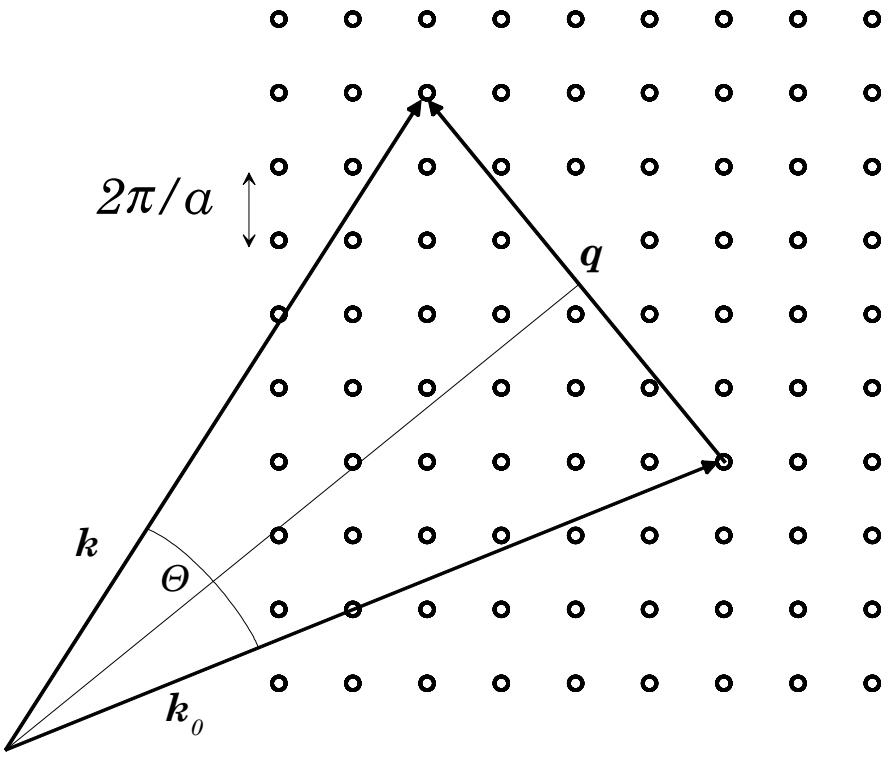

Figure 3.5: Géométrie de la diffusion par un cristal.

Il n'existe de solution en termes de Θ que si:

$$q < 2k \ . \tag{3.102}$$

La plus petite distance non nulle possible sur le réseau réciproque est sa maille, $2\pi/a$. On n'aura donc de solution dans une direction différente de la direction incidente que si $2\pi/a < 2k$, ce qu'on peut encore écrire:

$$a > \lambda/2 \ . \tag{3.103}$$

Si la longueur d'onde incidente est plus grande que 2a, il n'y a que la diffusion vers l'avant. Nous étendons donc la validité des résultats du paragraphe précédent, établis a priori pour une distance moyenne entre atomes très petite devant λ. Pour $\lambda < 2a$, il y a au moins une direction de diffusion efficace. Pour une longueur d'onde incidente suffisamment petite, on doit voir plusieurs directions correspondant à plusieurs points du réseau réciproque.

Cette diffusion cohérente dans des directions privilégiées est appelée diffusion de Bragg. Elle a en effet été largement utilisée par Bragg pour les premières déterminations de structures cristallines par diffusion de rayonnement X. C'est encore une méthode d'étude précieuse des matériaux. En enregistrant la lumière diffusée par un cristal, on enregistre les points du réseau réciproque. Il est alors facile de remonter, par transformation de Fourier, au réseau cristallin original. Nous avons considéré uniquement le cas d'atomes identiques. Dans un cristal plus complexe, on observe la superposition des figures de diffusion dues aux différentes espèces cristallines. Les différences d'amplitude de diffusion individuelles (dans le terme $\mathbf{e}_d\alpha$) permettent de remonter à l'arrangement tridimensionnel des atomes. L'ensemble de ces techniques peut être transposé à la diffusion d'autres types de rayonnement. Le fonctionnement ultime des microscopes électroniques correspond à une diffusion de Bragg des ondes électroniques sur le réseau cristallin. Les neutrons sont aussi largement utilisés. Ils interagissent en effet beaucoup mieux que les rayons X avec certains atomes légers.

3.3.4 Diffusion par un milieu désordonné

Considérons maintenant la diffusion par un milieu désordonné. Nous supposerons que les positions des centres diffuseurs sont aléatoires et que la densité n'est pas trop grande. Le nombre de centres diffuseurs dans un élément de volume de l'ordre de λ^3 est donc aléatoire. Nous pouvons comprendre à partir de cela les propriétés qualitatives du rayonnement diffusé.

$\mathbf{E}(\mathbf{q})$ est en effet la transformée de Fourier en $\mathbf{q}$ de δn. Si q est très voisin de zéro (diffusion vers l'avant), $\mathbf{E}$ est déterminé par une moyenne à grande échelle (extension spatiale de l'ordre de $1/q$) de la fluctuation d'indice. Même si le nombre de particules varie à l'échelle de λ, l'indice moyenné sur une échelle beaucoup plus grande est constant. On retrouve donc, vers l'avant, une diffusion cohérente de tous les atomes, comme dans le cas du milieu dense. Là encore, la notion d'indice de réfraction ordinaire prend tout son sens.

En revanche, si on s'intéresse au rayonnement diffusé dans une direction quelconque, on échantillonne dans δn les variations spatiales à une échelle de l'ordre de λ, qui sont importantes par hypothèse. Il y aura donc de la lumière diffusée. De plus, si δn est réellement aléatoire, il se comporte comme un bruit blanc dont la transformée de Fourier est pratiquement indépendante de $\mathbf{q}$. On peut donc s'attendre à une transformée de Fourier isotrope et à un diagramme de diffusion dominé par le terme $\mathbf{e}_d$ (il ne saurait y avoir de diffusion selon Ox quelle que soit la géométrie du milieu).

Précisons un peu ce raisonnement très qualitatif en nous intéressant à l'intensité diffusée dans une direction $\mathbf{k}$ très différente de la direction initiale. Cette intensité est proportionnelle à la composante de Fourier de la fonction d'autocorrélation de l'indice à une fréquence spatiale de l'ordre de λ. Pour estimer cette fonction d'autocorrélation, divisons le milieu en petits éléments de volume ΔV de taille de l'ordre de λ^3. Nous supposerons que le nombre d'atomes p dans chaque élément de volume est une fonction aléatoire et que les nombres de particules dans des volumes adjacents sont des fonctions aléatoires indépendantes (nous prenons ici une image statique du gaz, un instantané de la position de toutes les molécules dans un gaz réel. Cette approximation est donc raisonnable).

La fonction d'autocorrélation d'indice peut s'écrire, à un facteur $|\alpha|^2/4$ près, comme la fonction d'autocorrélation de la densité numérique $N(\mathbf{r})$ des particules:

$$\Gamma(\boldsymbol{\rho}) = \frac{|\alpha|^2}{4} \int N(\mathbf{r}) N(\mathbf{r} - \boldsymbol{\rho})\, d\mathbf{r} \ . \tag{3.104}$$

La densité numérique comporte un terme moyen, N_0, auquel s'ajoutent les fluctuations N_f: $N = N_0 + N_f$. La valeur moyenne spatiale de ces fluctuations est évidemment nulle. Si on reporte dans l'expression ci–dessus cette décomposition de N, on voit donc apparaître seulement deux termes. L'un est l'intégrale de N_0^2 et ne contribue qu'à la diffusion vers l'avant. C'est le seul terme qui intervient dans le cas du milieu homogène. On peut donc, pour la diffusion latérale, l'oublier. L'autre terme est l'autocorrélation de la fluctuation de densité. On peut donc écrire:

$$\Gamma(\boldsymbol{\rho}) = \frac{|\alpha|^2}{4} \int_V N_f(\mathbf{r}) N_f(\mathbf{r} - \boldsymbol{\rho})\, d\mathbf{r} \ . \tag{3.105}$$

Avec nos hypothèses, les fluctuations de densité de deux cellules adjacentes ne sont pas corrélées. Γ sera donc nulle dès que $\boldsymbol{\rho}$ est plus grand que la taille caractéristique d'une cellule, c'est à dire λ. Essentiellement, à l'échelle du volume total du milieu, Γ n'est non nulle que dans un voisinage de l'origine d'extension ΔV. Estimons maintenant l'ordre de grandeur de Γ dans ce voisinage, c'est à dire l'ordre de grandeur de $\Gamma(0)$. En supposant le milieu invariant par translation,

$$\Gamma(0) = \frac{|\alpha|^2}{4} V N_f^2 \tag{3.106}$$

Soit p le nombre de particules dans une cellule donnée. Il doit obéir à une loi de Poisson avec une valeur moyenne $N\Delta V$. L'écart quadratique moyen de p, Δp^2 doit donc être égal à cette valeur

moyenne. La densité numérique dans cette cellule est $p/\Delta V$. N_f^2 est l'écart quadratique moyen de la densité numérique dans les cellules. On a donc

$$N_f^2 \simeq \frac{\Delta p^2}{\Delta V^2} \simeq \frac{p}{\Delta V^2} = \frac{N}{\Delta V} \ . \tag{3.107}$$

L'ordre de grandeur de la fonction d'autocorrélation est donc:

$$\Gamma(0) = \frac{|\alpha|^2}{4} N \frac{V}{\Delta V} \tag{3.108}$$

Un modèle simple de Γ est donc de prendre une valeur constante égale à $\Gamma(0)$ dans le volume ΔV autour de l'origine et zéro partout ailleurs. La transformation de Fourier, en ordres de grandeur, se ramène alors à une multiplication par ΔV et l'intensité diffusée dans une direction loin de l'incidence est donc finalement proportionnelle à:

$$I \simeq |\mathbf{e}_d|^2 |\alpha|^2 N V \simeq |\mathbf{e}_d|^2 |\alpha|^2 \mathcal{N} \ . \tag{3.109}$$

L'intensité est ici seulement proportionnelle au nombre d'atomes, alors qu'elle était proportionnelle à son carré pour la diffusion par un milieu dense. C'est là un comportement caractéristique de l'émission incohérente, dont nous avions déjà rencontré un exemple avec la diffusion par un réseau aléatoire d'ouvertures. En fait, le rayonnement diffusé est simplement $\mathcal{N}$ fois le rayonnement diffusé par un atome unique (le diagramme de rayonnement étant donc essentiellement celui d'un dipôle unique). Ce résultat légitime les calculs de longueur d'extinction que nous avions faits pour la diffusion par l'atmosphère.

3.3.5 Influence de la dynamique du milieu

Dans beaucoup de cas, on ne peut pas considérer les atomes ou les centres diffuseurs comme immobiles. Si nos approximations sont réalistes pour un milieu dense ou un cristal (jusqu'à un certain point), elles tombent sûrement pour un gaz. Le rayonnement diffusé n'est plus monochromatique. On peut le comprendre qualitativement simplement en introduisant l'effet Doppler sur l'onde diffusée par une particule mobile. Pour décrire un tel champ et son spectre, on définit une densité spectrale d'énergie rayonnée, $J(\nu)$, telle que l'intensité dans une bande de fréquence $d\nu$ (on prendra garde que ν est en fait ici une pulsation) autour de ν soit $J(\nu)d\nu$. On peut alors définir la diffusion par la fonction $J(\mathbf{q}, \nu)$ qui donne la densité spectrale du rayonnement diffusé dans la direction définie par $\mathbf{q}$. On montre, par des arguments très similaires à ceux que nous avons utilisés, que cette fonction est la transformée de Fourier spatio–temporelle de la fonction d'autocorrélation spatio–temporelle des fluctuations d'indice (définies à partir de la polarisabilité) et qui dépendent cette fois de la position et du temps:

$$J(\mathbf{q}, \nu) = \frac{2}{\pi} |\mathbf{e}_d|^2 \int d\tau \, d\boldsymbol{\rho} \, e^{i\mathbf{q}\cdot\boldsymbol{\rho}} e^{-i(\nu-\omega)\tau} \Gamma(\boldsymbol{\rho}, \tau) \ , \tag{3.110}$$

avec

$$\Gamma(\boldsymbol{\rho}, \tau) = \int d\mathbf{r} \, \overline{\delta n^*(\mathbf{r}, t_0) \delta n(\mathbf{r} - \boldsymbol{\rho}, t_0 - \tau)} \ . \tag{3.111}$$

Dans la dernière expression, $\overline{\delta n}$ représente une moyenne temporelle sur t_0 à une échelle de temps longue par rapport aux temps du problème.

Pour un milieu statique, toutes les fonctions de ν se ramènent à un Dirac centré à la fréquence incidente ω et toutes les discussions précédentes restent valables. Pour un milieu quelconque, la fréquence est modifiée. Il est facile de voir que, pour un gaz, par exemple, la diffusion reste isotrope, mais que le spectre du rayonnement diffusé est essentiellement le spectre Doppler du milieu. Plus généralement, l'étendue spectrale du rayonnement doit être l'inverse du temps caractéristique de la fonction de corrélation (qui est presque toujours maximale à temps nul et décroissante ensuite). Pour

l'effet Doppler, le temps caractéristique est le temps nécessaire pour que les molécules se déplacent de λ, soit $\lambda/v = 2\pi c/\omega v$ où v est la vitesse caractéristique des molécules. On retrouve bien une largeur Doppler relative égale à v/c. Le calcul précis des fonctions de corrélation est souvent assez délicat et nous ne l'aborderons pas d'avantage ici.

Nous allons en revanche appliquer la formule précédente à un cas simple et intéressant technologiquement, la diffusion Brillouin, ou diffusion d'une onde lumineuse par une onde sonore. C'est, comme nous le verrons, un phénomène important pour comprendre les bruits dans les fibres optiques. C'est aussi le phénomène qui constitue les modulateurs acousto-optiques, très utilisés en physique des lasers.

Nous considérons donc un milieu parcouru par une onde acoustique sinusoïdale de fréquence ω_s. Cette onde correspond à une modulation de la densité du milieu et donc à une modulation de l'indice de réfraction. On peut donc écrire l'indice au point $\mathbf{r}$ et à l'instant t sous la forme:

$$n(\mathbf{r}, t) = n_0 + n_1 \cos(\mathbf{k}_s \cdot \mathbf{r} - \omega_s t) \ . \tag{3.112}$$

Dans cette expression, n_0 est l'indice moyen du milieu. Comme dans le paragraphe précédent, cet indice uniforme et statique ne contribue qu'à la propagation dans la direction initiale. Tous les phénomènes non triviaux sont contenus dans le second terme, dont les moyennes spatiales et temporelles sont nulles. Nous n'utiliserons donc, dans le calcul des fonctions de corrélation, que ce second terme.

Le vecteur d'onde $\mathbf{k}_s$ a pour module ω_s/v_s, où v_s est la vitesse du son dans le milieu. Pour fixer les ordres de grandeur, nous considérerons un milieu cristallin, où la vitesse du son est de quelques kilomètres par seconde: $v_s \simeq 10^3$ m/s. Nous considérerons une fréquence résolument ultrasonore, $\omega_s \simeq 6.10^8$ (soit 100 MHz en unités de fréquence). k_s est alors de l'ordre de 6.10^5/m, soit une longueur d'onde de l'ordre de $2\pi/k_s = 10$ μm. De manière évidente, il n'y aura d'effet sensible de l'onde sonore que si elle n'est pas trop grande par rapport à la longueur d'onde optique dans le matériau, c'est à dire si la fréquence acoustique est au moins de cet ordre. Pour des fréquences d'onde sonore très basses, on a simplement une propagation vers l'avant avec une petite modulation de phase.

Il est facile alors d'écrire la fonction de corrélation, en prenant, comme l'indique sa définition, l'onde acoustique sous forme complexe $n_1 \exp(i(\mathbf{k}_s \cdot \mathbf{r} - \omega_s t))$:

$$\Gamma(\boldsymbol{\rho}, \tau) = n_1^2 \int d\mathbf{r}\, \overline{e^{-i(\mathbf{k}_s \cdot \mathbf{r} - \omega_s t_0)} e^{i(\mathbf{k}_s \cdot (\mathbf{r} - \boldsymbol{\rho}) - \omega_s (t_0 - \tau))}} \ . \tag{3.113}$$

Le temps t_0 s'élimine naturellement et il ne reste qu'un quantité invariante dans le processus de moyennage temporel. La dépendance en $\mathbf{r}$ s'annulant aussi, l'intégrale de volume se ramène à une simple multiplication par V et on trouve finalement:

$$\Gamma(\boldsymbol{\rho}, \tau) = n_1^2 V e^{-i(\mathbf{k}_s \cdot \boldsymbol{\rho} - \omega_s \tau)} \ . \tag{3.114}$$

La densité spectrale de rayonnement diffusé s'exprime alors simplement. La transformée de Fourier spatiale de l'exponentielle complexe dans Γ se ramène, pour un milieu de grande extension, à une fonction de Dirac. La transformée de Fourier temporelle donne exactement une fonction de Dirac. On trouve donc finalement, à des coefficients numériques près qui contiennent diagramme de rayonnement des dipôles et des termes en n_1^2 proportionnels à l'intensité de l'onde sonore:

$$J(\mathbf{q}, \nu) \propto \delta(\mathbf{q} - \mathbf{k}_s) \delta(\nu - \omega_s - \omega) \mathcal{N}^2 \ . \tag{3.115}$$

Il ne sort donc du milieu, en plus de l'onde dans la direction initiale, qu'un seul faisceau diffusé, dans la direction définie par $\mathbf{q} + \mathbf{k}_s$. La fréquence de ce faisceau est décalée par rapport à la fréquence de l'onde incidente de ω_s. On peut comprendre assez facilement ce résultat en faisant appel à la notion de phonon. La propagation d'ondes acoustiques dans un réseau cristallin peut en effet se comprendre comme la propagation de quasi–particules, très analogues au photon, que l'on nomme phonons. Ces particules, d'énergie $\hbar\omega_s$, ont une quantité de mouvement $\hbar\mathbf{k}_s$. Les photons incidents ont une impulsion

$\hbar\mathbf{k}$ et une énergie $\hbar\omega_s$. La collision entre un photon et un phonon, si elle respecte la conservation de l'impulsion et de l'énergie, doit donc produire un photon d'impulsion $\hbar(\mathbf{k}+\mathbf{k}_s)$ et d'énergie $\hbar(\omega+\omega_s)$. Si ces arguments très simples ne permettent pas de donner la probabilité que l'événement se produise (c'est à dire de calculer l'intensité de l'onde diffusée), ils permettent néanmoins de comprendre les caractéristiques essentielles de l'onde produite.

Cette diffusion cohérente d'une onde lumineuse sur une onde sonore joue un grand rôle en optique. Un rôle néfaste, d'abord, pour les transmissions par fibre optique. Les collisions des photons avec les ondes sonores thermiquement excitées dans la fibre sont une cause importante de bruit de phase et de perte de signal optique. Il est très difficile de s'en affranchir. Les vibrations couplées efficacement aux ondes lumineuses ont une énergie de quelques centaines de MHz et sont donc largement peuplées par le bruit thermique de phonons.

En revanche, la diffusion Brillouin peut être mise à profit pour dévier et transposer en fréquence un faisceau laser. C'est le principe du modulateur acousto–optique, très utilisé maintenant. L'intérêt en est essentiellement la transposition de fréquence. A partir d'un seul faisceau laser, on peut disposer d'autres faisceaux de fréquences différentes mais ayant une relation de phase parfaitement bien définie avec le faisceau initial. L'exploration de fréquence que l'on peut atteindre dépend essentiellement de la bande passante des transducteurs électromécaniques créant l'onde. On peut maintenant couvrir une gamme de quelques centaines de MégaHertz. En optimisant le fonctionnement, l'efficacité de conversion entre la puissance incidente et la puissance transposée peut avoisiner 90%. On peut donc envisager de cascader ces modulateurs pour étendre l'excursion en fréquence. Enfin, en modulant temporellement l'intensité de l'onde sonore, on peut moduler l'intensité de l'onde diffusée sans changer sa direction ni sa fréquence. On peut ainsi réaliser des obturateurs très rapides, avec des temps de réponse très inférieurs à la microseconde. La déflexion du faisceau, enfin, peut être utilisée pour réaliser des balayages rapides. Les premières imprimantes laser fonctionnaient sur ce principe. Les modulateurs acousto-optiques, très coûteux, ont depuis été remplacés par des miroirs holographiques rotatifs.

Appendice 1

Réaction de rayonnement: modèle d'Abraham–Lorentz

Nous avons introduit la force de réaction de rayonnement par un simple argument de bilan énergétique. Nous allons, dans cet appendice, donner un modèle très naïf de particule chargée qui prédit correctement l'expression de cette force et qui montre qu'elle est en fait la résultante des forces que la particule exerce sur elle-même, non nulle si l'accélération de la particule varie. Nous montrerons aussi que ce modèle a des limites sérieuses (non–invariance relativiste, prédiction d'une masse électromagnétique incorrecte) et qu'il ne faut donc le prendre que comme un guide qualitatif.

Nous considérerons donc la particule chargée comme une distribution volumique de charges à symétrie sphérique. Si la particule était immobile, la distribution de charges s'écrirait simplement $\rho(\mathbf{r}) = \rho(r)$, l'origine étant bien sûr choisie au centre. Nous considérerons une particule en mouvement d'ensemble non relativiste. La distribution de charge se déplace donc en bloc avec une vitesse $\mathbf{v}(t)$. La densité de charges résultante s'écrit donc $\rho(\mathbf{r}, t)$ et la densité de courant associée $\mathbf{j}(\mathbf{r}, t) = \rho\mathbf{v}(t)$. Nous choisirons l'origine au centre de la particule à un instant t donné. La particule est placée dans un champ électromagnétique extérieur $\mathbf{E}_x, \mathbf{B}_x$. Nous noterons $\mathbf{E}_s$ et $\mathbf{B}_s$ les champs électriques et magnétiques produits par la particule en mouvement elle–même. Les champs totaux sont donc la somme de ces champs et des champs extérieurs.

Nous supposerons que toutes les propriétés énergétiques ou inertielles de la particule sont liées à son seul champ électromagnétique. La "masse mécanique" de la particule est donc nulle. L'équation du mouvement impose donc que la somme des forces de Lorentz dues au champ extérieur $\mathbf{F}_x$ et au champ propre $\mathbf{F}_s$ s'annulent. Nous devons donc avoir:

$$\mathbf{F}_x = -\int \rho\mathbf{E}_s + \mathbf{j} \times \mathbf{B}_s \ , \tag{1.1}$$

où l'intégrale s'étend à l'ensemble du volume de la particule. Nous allons exprimer cette intégrale en termes des dérivées successives de la vitesse de la particule. Nous aurons ainsi écrit une équation du mouvement décrivant l'évolution de la vitesse de la particule dans un champ extérieur. Nous verrons que le terme en dérivée seconde de la vitesse donne la force de réaction de rayonnement et que le terme en dérivée de la vitesse permet de définir une masse d'origine électromagnétique pour la particule.

Avant d'entrer dans les détails du calcul, il est important de préciser que ce modèle n'est pas très réaliste. Son principal défaut est d'imaginer une distribution de charges qui reste stable sous l'influence des seules forces électromagnétiques. Il est évident que notre répartition de charges devrait exploser littéralement sous l'influence des forces de Coulomb. En admettant même qu'on puisse décrire une particule de façon aussi naïve, il faudrait ajouter au modèle des forces de cohésion de nature non électromagnétique. A ces forces serait associée une énergie et donc une contribution à l'inertie. Le calcul est faisable et permet de corriger certains défauts du modèle mais la description en termes de distribution étendue des particules élémentaires reste bien naïve.

Nous devons donc calculer l'intégrale des forces de Lorentz produites par la particule elle–même. Il est tout à fait évident d'abord que, pour un mouvement non relativiste, les forces magnétiques jouent un rôle négligeable. Nous écrirons donc:

$$\mathbf{F}_x = -\int \rho(\mathbf{r},t)\mathbf{E}_s(\mathbf{r},t)\, d^3\mathbf{r} \ . \tag{1.2}$$

Le champ $\mathbf{E}_s$ dérive des potentiels retardés V_s et $\mathbf{A}_s$. On a:

$$V_s(\mathbf{r},t) = \frac{1}{4\pi\epsilon_0}\int \frac{\rho(\mathbf{r}',t-R/c)}{R}\, d^3\mathbf{r}' \tag{1.3}$$

$$\mathbf{A}_s(\mathbf{r},t) = \frac{1}{4\pi\epsilon_0 c^2}\int \frac{\mathbf{j}(\mathbf{r}',t-R/c)}{R}\, d^3\mathbf{r}' \ , \tag{1.4}$$

où $R = |\mathbf{R}|$ avec $\mathbf{R} = \mathbf{r} - \mathbf{r}'$. L'extension de la particule étant petite, les temps retardés sont pratiquement égaux à t. On peut donc effectuer un développement de Taylor à tous les ordres des potentiels en puissance du retard R/c.

Pour le potentiel vecteur, on a:

$$\mathbf{A}_s(\mathbf{r},t) = \frac{1}{4\pi\epsilon_0 c^2}\sum_{n=0}^{\infty}\frac{(-1)^n}{n!c^n}\int d^3\mathbf{r}'\, R^{n-1}\frac{\partial^n}{\partial t^n}\mathbf{j}(\mathbf{r}',t) \ . \tag{1.5}$$

On peut alors facilement exprimer $-\partial\mathbf{A}_s/\partial t$ et donner la contribution du champ électromoteur à l'intégrale de la force de Lorentz:

$$\mathbf{F}_{x,\mathbf{A}} = \frac{1}{4\pi\epsilon_0 c^2}\sum\frac{(-1)^n}{n!c^n}\int d^3\mathbf{r}\, d^3\mathbf{r}'\, \rho(\mathbf{r},t)R^{n-1}\frac{\partial^{n+1}}{\partial t^{n+1}}\mathbf{j}(\mathbf{r}',t) \ . \tag{1.6}$$

Notons que dans la suite du calcul, les sommes porteront implicitement sur n, entre zéro et l'infini.

Pour le potentiel scalaire le même développement donne:

$$V_s(\mathbf{r},t) = \frac{1}{4\pi\epsilon_0}\sum\frac{(-1)^n}{n!c^n}\int d^3\mathbf{r}'\, R^{n-1}\frac{\partial^n}{\partial t^n}\rho(\mathbf{r}',t) \ , \tag{1.7}$$

dont on calcule facilement le gradient par rapport à $\mathbf{r}$ (notons que cette dérivation commute avec l'intégrale sur $\mathbf{r}'$). On trouve finalement une contribution à la force de Lorentz:

$$\mathbf{F}_{x,V} = \frac{1}{4\pi\epsilon_0}\sum\frac{(-1)^n}{n!c^n}\int d^3\mathbf{r}\, d^3\mathbf{r}'\, \rho(\mathbf{r},t)\boldsymbol{\nabla}_{\mathbf{r}}R^{n-1}\frac{\partial^n}{\partial t^n}\rho(\mathbf{r}',t) \ . \tag{1.8}$$

Considérons les deux premiers termes dans le développement de $\mathbf{F}_{x,V}$. Le premier fait intervenir:

$$\int d^3\mathbf{r}\, d^3\mathbf{r}'\, \rho(\mathbf{r},t)\rho(\mathbf{r}',t)\boldsymbol{\nabla}_{\mathbf{r}}R \ . \tag{1.9}$$

On y reconnaît facilement un terme proportionnel à l'intégrale sur la distribution de charges de la force de Coulomb exercée par les autres parties de la distribution. Le principe de l'action et de la réaction, vérifié par les forces électrostatiques, montre alors que cette intégrale est nulle. Le terme d'ordre un est nul lui aussi, puisqu'il fait intervenir l'intégrale du gradient d'une constante. Le premier terme non nul est donc au moins à l'ordre 2. On peut en fait regrouper, au prix d'une simple renumérotation, les termes d'ordre n issus du potentiel vecteur et les termes d'ordre $n+2$ issus du potentiel scalaire et écrire finalement:

$$\mathbf{F}_x = \frac{1}{4\pi\epsilon_0 c^2}\sum\frac{(-1)^n}{n!c^n}\int d^3\mathbf{r}\, \rho(\mathbf{r},t)\frac{\partial^{n+1}}{\partial t^{n+1}}\mathbf{C}_n(\mathbf{r},t) \ , \tag{1.10}$$

avec:

$$\mathbf{C}_n(\mathbf{r},t) = \int d^3\mathbf{r}'\, R^{n-1}\left[\mathbf{j}(\mathbf{r}',t) + \frac{\partial\rho(\mathbf{r}',t)}{\partial t}\frac{\boldsymbol{\nabla}_{\mathbf{r}}R^{n+1}}{(n+1)(n+2)R^{n-1}}\right] \ . \tag{1.11}$$

La contribution du deuxième terme du crochet dans $\mathbf{C}_n$ peut être transformée en utilisant l'équation de conservation de la charge: $\partial\rho/\partial t = -\boldsymbol{\nabla}_{\mathbf{r}'} \cdot \mathbf{j}$. On remarquera aussi que:

$$\boldsymbol{\nabla}_{\mathbf{r}} R^{n+1} = (n+1)\mathbf{R}R^{n-1} \ . \tag{1.12}$$

Le deuxième terme du crochet se réécrit donc:

$$-\frac{1}{n+2}\int d^3\mathbf{r}'\, \boldsymbol{\nabla}_{\mathbf{r}'} \cdot \mathbf{j}(\mathbf{r}',t) R^{n-1}\mathbf{R} \ . \tag{1.13}$$

En utilisant alors une formule peu connue d'intégration par parties:

$$\int \mathbf{a}(\boldsymbol{\nabla}\cdot\mathbf{b}) = -\int (\mathbf{b}\cdot\boldsymbol{\nabla})\mathbf{a} \ , \tag{1.14}$$

valable si les champs s'annulent à l'infini, ce qui est évident ici, on peut écrire ce deuxième terme sous la forme:

$$\frac{1}{n+2}\int d^3\mathbf{r}'\, (\mathbf{j}\cdot\boldsymbol{\nabla}_{\mathbf{r}'}) R^{n-1}\mathbf{R} \ . \tag{1.15}$$

On a alors:

$$(\mathbf{j}\cdot\boldsymbol{\nabla}_{\mathbf{r}'})R^{n-1}\mathbf{R} = R^{n-1}(\mathbf{j}\cdot\boldsymbol{\nabla}_{\mathbf{r}'})\mathbf{R} + [(\mathbf{j}\cdot\boldsymbol{\nabla}_{\mathbf{r}'})R^{n-1}]\mathbf{R} \ , \tag{1.16}$$

et

$$(\mathbf{j}\cdot\boldsymbol{\nabla}_{\mathbf{r}'})R^{n-1} = -(n-1)R^{n-1}\frac{\mathbf{j}\cdot\mathbf{R}}{R^2}\mathbf{R} \ , \tag{1.17}$$

ainsi que:

$$(\mathbf{j}\cdot\boldsymbol{\nabla}_{\mathbf{r}'})\mathbf{R} = -\mathbf{j} \ . \tag{1.18}$$

En regroupant tous ces termes, on met finalement l'intégrale $\mathbf{C}_n$ sous la forme:

$$\mathbf{C}_n(\mathbf{r},t) = \int d^3\mathbf{r}'\, R^{n-1}\left[\frac{n+1}{n+2}\rho(\mathbf{r}',t)\mathbf{v}(t) - \frac{n-1}{n+2}\rho(\mathbf{r}',t)\frac{\mathbf{v}(t)\cdot\mathbf{R}}{R^2\mathbf{R}}\right] \ , \tag{1.19}$$

où nous avons utilisé $\mathbf{j} = \rho\mathbf{v}$. Pour des raisons de symétrie évidente, cette intégrale doit avoir la direction de la vitesse $\mathbf{v}$. Dans l'intégrale, toutes les orientations de $\mathbf{R}$ par rapport à $\mathbf{v}$ apparaissent de façon symétrique. Il est donc possible, sans changer le résultat final, de remplacer le deuxième terme du crochet par la valeur moyenne de sa projection sur $\mathbf{v}$ pour toutes les orientations possibles de $\mathbf{R}$. Cette valeur est proportionnelle à la moyenne du cosinus carré de l'angle entre $\mathbf{v}$ et $\mathbf{R}$, soit $1/3$. On a donc:

$$\mathbf{C}_n(\mathbf{r},t) = \frac{2}{3}\mathbf{v}(t)\int d^3\mathbf{r}'\, R^{n-1}\rho(\mathbf{r}',t) \ , \tag{1.20}$$

un résultat étonnamment simple.

Reportons maintenant cette expression dans celle de la force de Lorentz. Nous devons prendre la $n+1^{\text{ème}}$ dérivée temporelle de $\mathbf{C}_n$. Il apparaît ainsi les termes en dérivées successives de la vitesse. Mais il apparaît aussi des termes faisant intervenir les dérivées temporelles de ρ. $\mathbf{v}$ étant la vitesse d'ensemble, $\partial\rho/\partial t \simeq (\boldsymbol{\nabla}\rho)\cdot\mathbf{v}$. Les termes en dérivées de ρ sont donc des termes quadratiques dans la vitesse et ses dérivées. Nous pourrons les négliger si la vitesse est suffisamment faible et écrire finalement:

$$\mathbf{F}_x = \frac{1}{4\pi\epsilon_0 c^2}\sum \frac{(-1)^n}{n!c^n}\frac{\partial^{n+1}\mathbf{v}(t)}{\partial t^{n+1}}\frac{2}{3}\int d^3\mathbf{r}\, d^3\mathbf{r}'\, \rho(\mathbf{r},t)\rho(\mathbf{r}',t)R^{n-1} \ . \tag{1.21}$$

Explicitons le terme d'ordre 0:

$$\mathbf{F}_x^0 = \frac{4}{3}\frac{U}{c^2}\dot{\mathbf{v}} \ , \tag{1.22}$$

où

$$U = \frac{1}{4\pi\epsilon_0}\int d^3\mathbf{r}\, d^3\mathbf{r}'\, \frac{\rho(\mathbf{r},t)\rho(\mathbf{r}',t)}{R} \ , \tag{1.23}$$

est simplement l'énergie électrostatique de la distribution de charges. A l'ordre 1, nous avons:

$$\mathbf{F}_x^1 = -\frac{1}{6\pi\epsilon_0}\frac{q^2}{c^3}\ddot{\mathbf{v}} \ , \tag{1.24}$$

où

$$q = \int d^3\mathbf{r}\,\rho(\mathbf{r},t) \ , \tag{1.25}$$

est simplement la charge totale de la distribution. On peut donc écrire finalement, en négligeant les termes d'ordre supérieur, l'équation du mouvement de la particule sous la forme:

$$m\dot{\mathbf{v}} = \mathbf{F}_x + \mathbf{F}_r \ , \tag{1.26}$$

où $\mathbf{F}_x$ est la résultante des forces extérieures et $\mathbf{F}_r$ la force de réaction de rayonnent que nous obtenons ici directement sous la forme démontrée par les arguments énergétiques dans le corps du chapitre. m est la "masse électromagnétique"

$$m = \frac{4}{3}\frac{U}{c^2} \ , \tag{1.27}$$

de la particule décrivant l'inertie de son propre champ.

Si ce calcul nous donne bien la force de réaction de rayonnement, il n'est guère satisfaisant pour la masse électromagnétique. Celle-ci est en effet égale à $4U/3c^2$ où U est l'énergie. L'invariance relativiste imposerait une masse égale à U/c^2. La différence est due bien sûr au caractère irréaliste de ce modèle qui ne tient pas compte des forces de cohésion nécessaires à la stabilité de la répartition de charges. La conséquence en est que ce modèle n'est pas invariant de Lorentz ce qui limite beaucoup son intérêt.

Partie V

Electromagnétisme dans la matière

Introduction

Nous nous proposons dans cette partie de décrire la propagation des champs électromagnétiques dans la matière dense. En fait, l'objet essentiel de ce chapitre sera de donner un contenu physique précis à la notion d'indice de réfraction et d'étudier la propagation de la lumière dans un certain nombre de milieux choisis pour leur caractère exemplaire.

Les équations de Maxwell doivent décrire de façon parfaitement satisfaisante la propagation dans les milieux matériels à condition d'inclure *toutes* les sources du champ électromagnétique. Il y a a priori deux sortes de courants ou de répartitions de charges. Il y a, d'une part, les courants macroscopiques produits par les générateurs ou les sources, en général contrôlés par l'expérimentateur, et descriptibles dans les termes que nous avons déjà employés dans les chapitres précédents. Il y a, d'autre part, tous les courants, plus ou moins incontrôlés, qui décrivent tous les transports de charge dans la matière dense. Tous les électrons d'un milieu peuvent contribuer à la propagation d'une onde. Nous appellerons charges (ou courants) *libres* les premiers, charges ou courants *liés* les seconds. Libres, parce qu'ils sont sous contrôle de l'expérimentateur, liés parce qu'ils correspondent aux mouvements de charges liées à la matière.

Il est évident que les charges libres ne nous poseront pas plus de problèmes que dans l'espace libre et qu'une adaptation des solutions en potentiels retardés ou des développements multipolaires suffira à les traiter. Ce qui rend le problème complexe, c'est bien sûr la présence des charges liées. On doit a priori tenir compte de toutes les charges du milieu et décrire les champs électriques intra–atomiques autant que celui de l'onde plane que nous essaierons de propager! On est en fait dans une situation analogue à celle de la mécanique quand elle essaie de décrire un gaz en tenant comptabilité de la position et de la vitesse de toutes les molécules. Comme en thermodynamique, nous n'aurons pas besoin de connaître explicitement les champs microscopiques dans leur horrible complexité. Il serait d'ailleurs tout à fait illusoire de chercher à les mesurer. Les seules quantités qui nous intéresseront seront des moyennes d'ensemble à l'échelle macroscopique de nos sources et de nos détecteurs.

L'objet du premier chapitre de cette partie sera donc de définir des processus de moyennage, de lissage, permettant de dériver des champs à l'échelle microscopique à partir des champs ayant une échelle de variation macroscopique. Cette échelle devra cependant être assez fine pour une description précise des phénomènes macroscopiques. Elle devra par exemple, pour un problème de propagation, être petite devant la longueur d'onde. Chaque atome, chaque molécule du milieu ne joue, dans ces moyennes, de rôle qu'à une distance très grande par rapport à sa taille. Nous pourrons donc utiliser les techniques de développements multipolaires du chapitre précédent pour traiter ce champ. Nous définirons donc des densités macroscopiques de polarisation, dipolaire ou quadripolaire électrique et dipolaire magnétique. Nous montrerons enfin que les équations décrivant les champs moyennés sont les équations de Maxwell, à condition d'y faire intervenir des densités de charge et de courant, elles aussi macroscopiques, liées aux dérivées spatiales et temporelles des densités de polarisation. Nous aurions donc résolu le problème si nous connaissions ces densités de polarisation.

En général, on ne peut rien dire des densités de polarisation. Elles dépendent en effet évidemment des champs appliqués qui, en retour, dépendent des densités de polarisation. Nous trouverons dans le deuxième chapitre une issue à ce problème. Nous supposerons, comme c'est pratiquement toujours le cas, que les champs appliqués seront assez faibles pour que le matériau réponde de façon

linéaire à l'excitation. Autrement dit, les densités de polarisation seront des fonctions linéaires des champs électriques et magnétiques macroscopiques, ce qui nous permettra enfin d'écrire les équations de propagation, les équations de Maxwell dans la matière, sous une forme close. Nous introduirons les susceptibilités électriques et magnétiques qui décrivent ces relations de proportionnalité, et nous donnerons quelques modèles simples de thermodynamique statistique pour prévoir leurs ordres de grandeur. Enfin, nous établirons sur ces susceptibilités, des résultats très généraux s'appliquant à priori à toute théorie de la réponse linéaire. Nous montrerons en effet que la simple causalité impose des relations très fortes entre les parties réelles et imaginaires des susceptibilités. Comme les unes décrivent l'indice de réfraction et les autres l'absorption par le milieu, nous en déduirons des relations très profonde entre dispersion et absorption.

Dans le dernier chapitre, enfin, nous appliquerons tout cela à des problèmes de propagation. Nous essaierons de comprendre pourquoi la plupart des milieux sont absorbants. Nous verrons à quelles conditions un milieu matériel peut être amplificateur pour les ondes lumineuses et nous inventerons en passant le laser. Nous envisagerons aussi la propagation dans les plasmas ou les métaux, équivalents de ce point de vue. Nous nous pencherons aussi sur les conditions de passage entre deux milieux différents. Nous en déduirons, dans le cas des diélectriques transparents, les lois de Descartes–Snell de la réfraction, mais aussi les coefficients de Fresnel qui décrivent les amplitudes relatives des ondes réfléchies et réfractées.

Chapitre 1

Equations de Maxwell dans la matière

La première étape de notre travail est donc d'écrire, de manière consistante, des valeurs moyennes macroscopiques à partir des champs et sources réels.

1.1 Champs et charges microscopiques et macroscopiques

Nous noterons ρ_t et $\mathbf{j}_t$ les densités de charges et de courant complètes, incluant les charges liées et les charges libres. Si nous notons $\mathbf{e}$ et $\mathbf{b}$ les vrais champs électrique et magnétique à l'échelle microscopique, ils sont déterminés par ρ_t et $\mathbf{j}_t$ et par les équations de Maxwell sous leur forme habituelle, qu'il n'est peut être pas nécessaire de rappeler encore une fois.

Nous voulons définir des moyennes des quantités ci dessus à une échelle très grande par rapport à la taille atomique, ou par rapport aux distances entre molécules voisines. Nous devrons d'autre part pouvoir décrire des phénomènes ondulatoires. Il faudra donc que l'échelle de moyennage reste petite par rapport à la longueur d'onde du rayonnement considéré. Nous choisirons donc l'échelle de notre moyennage, s_0, de telle manière que:

$$a_0 \ll s_0 \ll \lambda \ , \tag{1.1}$$

où a_0 est le rayon de Bohr, qui est aussi l'ordre de grandeur de la distance entre particules dans la matière dense. Si nous nous intéressons à des ondes optiques, les longueurs d'onde étant de l'ordre du micron et a_0 de la fraction d'Ångström, nous pouvons trouver facilement une échelle intermédiaire aux environs de la dizaine de nanomètres. Une sphère de rayon s_0 contient un très grand nombre d'atomes, tout en restant un petit objet à l'échelle de la propagation. Bien sûr, nous nous interdisons ainsi de décrire la propagation d'une onde de longueur d'onde comparable à a_0, dans le domaine des rayons X. La propagation d'une telle onde est beaucoup mieux décrite en termes de diffraction par le réseau cristallin qu'en termes d'indice de réfraction. Nous nous interdisons aussi de décrire des milieux très dilués. Mais il est évident que ces milieux sont essentiellement identiques au vide.

Nous définissons une fonction de lissage, d'extension s_0, $W(\mathbf{s})$, homogène à l'inverse d'un volume. Nous pourrons par exemple prendre une fonction gaussienne, de la forme $\exp(-s^2/s_0^2)/s_0^3$, avec un coefficient de normalisation convenable pour que l'intégrale sur tout l'espace de W soit égale à un. Notons que W doit être une fonction à symétrie sphérique pour respecter l'invariance par rotation des équations de Maxwell. Il est aussi important qu'elle varie régulièrement à l'échelle de s_0. Si nous prenions par exemple pour W la fonction caractéristique d'une sphère de rayon r_0, nous risquerions de ne pas obtenir des champs lisses à l'échelle macroscopique. Il suffirait en effet de déplacer cette sphère d'une quantité voisine de a_0 pour englober dans la moyenne un électron de plus ou de moins. Comme la matière est en général globalement neutre, on pourrait ainsi modifier considérablement la densité moyennée de charges par un déplacement microscopique de la sphère d'intégration. Notons également que les résultats de notre travail seront complètement indépendants de la forme de W et de l'échelle de moyennage s_0. Nous n'aurons d'ailleurs jamais à réaliser effectivement un de ces moyennages,

puisque nous ne connaîtrons jamais les champs microscopiques. Nous définissons alors les quantités macroscopiques moyennées comme:

$$\mathbf{E}(\mathbf{r},t) = \int \mathbf{e}(\mathbf{r}-\mathbf{s},t)W(\mathbf{s})\,d^3\mathbf{s} \tag{1.2}$$

$$\mathbf{B}(\mathbf{r},t) = \int \mathbf{b}(\mathbf{r}-\mathbf{s},t)W(\mathbf{s})\,d^3\mathbf{s} \tag{1.3}$$

$$\varrho(\mathbf{r},t) = \int \rho_t(\mathbf{r}-\mathbf{s},t)W(\mathbf{s})\,d^3\mathbf{s} \tag{1.4}$$

$$\mathbf{J}(\mathbf{r},t) = \int \mathbf{j}_t(\mathbf{r}-\mathbf{s},t)W(\mathbf{s})\,d^3\mathbf{s}\ . \tag{1.5}$$

Ces quantités ne peuvent, par construction, varier plus rapidement que s_0. Elles ont donc toutes les bonnes caractéristiques pour jouer le rôle de variables macroscopiques. Les équations de Maxwell étant linéaires, elles commutent avec cette opération de prise de valeur moyenne par une intégration. Les quantités moyennées obéissent donc rigoureusement aux équations de Maxwell.

Dans les sources, nous pourrons, comme nous le mentionnions dans l'introduction, séparer les contributions des charges libres et des charges liées. Nous écrirons ϱ et $\mathbf{J}$ sous la forme:

$$\varrho = \rho + \rho_\ell \tag{1.6}$$

$$\mathbf{J} = \mathbf{j} + \mathbf{j}_\ell\ , \tag{1.7}$$

où ρ et $\mathbf{j}$ sont les densités de charges libres, et ρ_ℓ, $\mathbf{j}_\ell$ les densités de charges et de courant liées à la matière. Nous ne nous préoccuperons plus pendant un moment des densités libres, qui se traitent comme dans le vide. Notons toutefois qu'il s'agit de quantités moyennées à l'échelle s_0. Il n'y aurait plus de sens à parler dans notre approche d'une charge ponctuelle ou de tenter de déterminer les champs sur des échelles de distance inférieures à s_0.

1.2 Distributions moyennées

Nous allons maintenant tenter de décrire les charges liées. Nous allons pour cela regrouper les charges de la matière en molécules. Cette approche est bien sûr justifiée par la composition de la matière, mais elle nous permettra, surtout, de décrire les champs produits par une molécule en termes de développements multipolaires.

1.2.1 Densités microscopiques

Nous diviserons donc l'énorme collection de charges liées en molécules. Une charge liée restera toujours attachée à la même molécule. Nous indicerons les molécules du milieu avec un indice m. Une molécule pouvant contenir plusieurs charges, nous indicerons les charges à l'intérieur de chaque molécule par un indice grec β. Nous noterons donc $q_{m\beta}$ la charge d'une des charges de la molécule m, $\mathbf{r}_{m\beta}$ sa position et $\mathbf{v}_{m\beta}$ sa vitesse. Les densités de charge et de courant liées avant moyennage sont données respectivement par:

$$\sum_{m,\beta} q_{m\beta}\delta(\mathbf{r}-\mathbf{r}_{m\beta}) \tag{1.8}$$

$$\sum_{m,\beta} q_{m\beta}\mathbf{v}_{m\beta}\delta(\mathbf{r}-\mathbf{r}_{m\beta})\ . \tag{1.9}$$

En plus de ces quelques notations, nous allons faire des hypothèses réalistes sur nos molécules. Nous allons tout d'abord les supposer immobiles. Tout au moins, nous supposerons qu'elles se déplacent suffisamment peu ou suffisamment lentement pour que ce déplacement n'affecte pas la propagation.

Comme nous ne manipulerons que des quantités moyennées à l'échelle s_0, nous n'aurons à exclure que les mouvements d'ensemble à une échelle macroscopique. L'agitation dans les gaz, par exemple, a un libre parcours moyen si petit (quelques dizaines d'Ångstöms à pression ambiante) que ces mouvements n'influent en rien sur la propagation d'une onde optique. Si nous avions un déplacement d'ensemble suffisamment rapide, la propagation en serait affectée. Un champ en un endroit crée une polarisation moyenne qui pourrait se propager en un autre endroit et contribuer au rayonnement d'une autre région: il apparaîtrait des termes de convection dans les équations de propagation. Notons toutefois qu'il serait possible de tenir compte de ces déplacements d'ensemble au prix de quelques complications arithmétiques sans importance. Nous expliquerons brièvement comment les équations de Maxwell dans la matière sont modifiées dans un milieu en mouvement. Nous négligerons aussi toute action mécanique du champ sur les particules. Nous ne considérerons pas le cas où le champ pourrait induire des déplacements d'ensemble ou des variations de densité. Très généralement, l'action des champs sur les degrés de liberté externes des molécules est complètement négligeable devant l'agitation thermique ou les forces de liaison des cristaux. Ce n'est qu'au voisinage immédiat d'une transition de phase ou d'un point critique, par exemple, que cette petite perturbation peut avoir un effet important sur le milieu. Il faut citer aussi les milieux dilués d'atomes refroidis par laser où on pourrait observer ce genre d'effets.

Si nous notons $\mathbf{r}_m$ la position de la m-ième molécule (par exemple la position de son centre de gravité, ou de tout autre point remarquable situé au sein de la molécule), nous aurons donc:

$$\dot{\mathbf{r}}_m = \mathbf{v}_m = 0 \ . \tag{1.10}$$

Nous allons supposer aussi que nos molécules sont neutres ($\sum_\rho q_{m\beta} = 0$). Si ce n'était pas le cas, elles ne pourraient créer, puisqu'elles sont statistiquement immobiles, qu'une répartition de charges électrostatiques, dont le champ serait aisément calculable et qui, de toutes façons, ne contribueraient aucunement au rayonnement.

Nous pouvons maintenant repérer la position des charges de la molécule m par rapport à la position constante de celle-ci. Nous écrirons donc:

$$\mathbf{r}_{m\beta} = \mathbf{r}_m + \boldsymbol{\xi}_{m\beta} \ , \tag{1.11}$$

où $\boldsymbol{\xi}$ est par hypothèse une quantité de l'ordre de a_0, taille de la molécule. $\mathbf{r}_m$ étant constant, on peut réécrire les densités de charge et de courant liées avant moyennage comme:

$$\sum_{m,\beta} q_{m\beta}\delta(\mathbf{r} - \mathbf{r}_m - \boldsymbol{\xi}_{m\beta}) \tag{1.12}$$

$$\sum_{m,\beta} q_{m\beta}\dot{\boldsymbol{\xi}}_{m\beta}\delta(\mathbf{r} - \mathbf{r}_m - \boldsymbol{\xi}_{m\beta}) \ . \tag{1.13}$$

1.2.2 Moments multipolaires

Après moyennage, on ne regarde les champs qu'à une échelle de l'ordre de s_0. La taille des molécules étant de l'ordre de a_0, il est naturel de traiter leurs champs par un développement multipolaire. Comme elles sont neutres, ne peuvent intervenir que les termes dipolaires électriques, à l'ordre dominant, et dipolaires magnétiques et quadripolaires électriques à l'ordre suivant. Nous nous arrêterons là. En général, les termes quadripolaires électriques jouent un rôle complètement négligeable pour la propagation dans la matière mais, comme ils apparaissent au même ordre de développement que les termes dipolaires magnétiques, il nous faut les conserver un temps pour assurer la cohérence de nos développements. Nous décrirons donc une molécule par son dipôle électrique:

$$\mathbf{p}_m(t) = \sum_\beta q_{m\beta}\boldsymbol{\xi}_{m\beta} \ , \tag{1.14}$$

son dipôle magnétique:

$$\mathbf{m}_m(t) = \frac{1}{2}\sum_\beta q_{m\beta}\boldsymbol{\xi}_{m\beta} \times \dot{\boldsymbol{\xi}}_{m\beta} \ , \tag{1.15}$$

et enfin son quadripôle électrique, dont les composantes s'écrivent:

$$q_{mij}(t) = \sum_\beta q_{m\beta}\xi_{m\beta i}\xi_{m\beta j} \ . \tag{1.16}$$

L'ensemble des multipôles ainsi recensés peuvent être décrits par des densités de polarisation, en ajoutant juste des fonctions δ centrées sur les molécules. Les densités de polarisation microscopiques sont donc données par:

$$\mathbf{p}(\mathbf{r},t) = \sum_m \mathbf{p}_m(t)\delta(\mathbf{r}-\mathbf{r}_m) \tag{1.17}$$

$$\mathbf{m}(\mathbf{r},t) = \sum_m \mathbf{m}_m(t)\delta(\mathbf{r}-\mathbf{r}_m) \tag{1.18}$$

$$\mathbf{q}(\mathbf{r},t) = \sum_m \mathbf{q}_m(t)\delta(\mathbf{r}-\mathbf{r}_m) \ . \tag{1.19}$$

Nous allons maintenant procéder au lissage de ces distributions microscopiques.

1.2.3 Densités macroscopiques

Le moyennage par la fonction W des densités microscopiques de charges liées redonne les charges liées qui peuvent donc s'écrire:

$$\rho_\ell(\mathbf{r},t) = \sum_{m,\beta} q_{m\beta}W(\mathbf{r}-\mathbf{r}_m-\boldsymbol{\xi}_{m\beta}) \tag{1.20}$$

$$\mathbf{j}_\ell(\mathbf{r},t) = \sum_{m,\beta} q_{m\beta}\dot{\boldsymbol{\xi}}_{m\beta}W(\mathbf{r}-\mathbf{r}_m-\boldsymbol{\xi}_{m\beta}) \ . \tag{1.21}$$

Il est en effet tout à fait évident que la convolution de la fonction de lissage W par une fonction de Dirac donne une fonction W du même argument que la fonction de Dirac initiale. On lit directement sur les expressions précédentes que l'échelle de variation de ces quantités ne peut pas être plus petite que s_0.

On définira de même des densités de polarisation moyennes en moyennant les quantités microscopiques introduites à la fin du paragraphe précédent. On aura:

$$\mathbf{P}(\mathbf{r},t) = \sum_m \mathbf{p}_m W(\mathbf{r}-\mathbf{r}_m) \tag{1.22}$$

$$\boldsymbol{\mathcal{M}}(\mathbf{r},t) = \sum_m \mathbf{m}_m W(\mathbf{r}-\mathbf{r}_m) \tag{1.23}$$

$$\boldsymbol{\mathcal{Q}}(\mathbf{r},t) = \sum_m \mathbf{q}_m W(\mathbf{r}-\mathbf{r}_m) \ . \tag{1.24}$$

On notera que $\boldsymbol{\mathcal{Q}}$, comme le moment quadripolaire d'une molécule unique, est, avec nos notations, un tenseur symétrique.

Ces densités de polarisation ont une signification physique bien plus profonde que les charges liées elles–même. Si rien ne permet à priori de mesurer les densités moyennes de charges liées, la densité de polarisation exprime que chaque élément de volume se comporte comme un petit dipôle électrique ou magnétique, dont le champ est au moins en principe mesurable. Il serait donc préférable de faire intervenir dans les équations de Maxwell macroscopiques les densités de polarisation plutôt que les charges liées, plus obscures. Nous allons maintenant montrer qu'on peut effectivement exprimer charges et courant liés en fonction des dérivées spatiales et temporelles des densités de polarisation. En quelque sorte, nous allons inverser la définition de ces densités de polarisation.

1.2.4 Expression des densités de charges liées

Densité de charges

Nous allons commencer par la densité de charges, donnée par (1.20). Nous allons développer, dans cette expression, la fonction W en puissance de $\boldsymbol{\xi}$. En effet, la fonction ne varie qu'à l'échelle s_0 et $\boldsymbol{\xi}$ est d'ordre a_0. On peut donc sans problèmes effectuer un développement de Taylor de chacun des termes de la source par rapport à $\boldsymbol{\xi}_{m\beta}$. On écrira:

$$W(\mathbf{r}-\mathbf{r}_m-\boldsymbol{\xi}_{m\beta}) = W(\mathbf{r}-\mathbf{r}_m) - \boldsymbol{\xi}_{m\beta}\cdot\boldsymbol{\nabla}W(\mathbf{r}-\mathbf{r}_m) + \frac{1}{2}\sum_{i,j}\xi_{m\beta i}\xi_{m\beta j}\partial_i\partial_j W \ , \tag{1.25}$$

où les dérivées de W s'entendent par rapport à son argument $\mathbf{r}$ et doivent toutes être évaluées en $\mathbf{r}-\mathbf{r}_m$. Nous avons poussé ce développement à l'ordre 2. Nous verrons que c'est nécessaire pour retrouver les termes quadripolaires électriques. En portant ce développement dans (1.20), on trouve la densité de charges sous la forme d'une somme de trois termes:

$$\rho_\ell = \rho_\ell^{(0)} + \rho_\ell^{(1)} + \rho_\ell^{(2)} \ , \tag{1.26}$$

faisant intervenir les puissances successives de ξ.

Il est évident que le terme d'ordre 0 s'annule. La fonction W, constante, sort de la somme sur les charges de la molécule m et il reste un terme en $\sum_\beta q_{m\beta}$, nul puisque la molécule est neutre. Le terme d'ordre 1 s'écrit:

$$\rho_\ell^{(1)} = -\sum_m \boldsymbol{\nabla}W(\mathbf{r}-\mathbf{r}_m)\cdot\sum_\beta q_{m\beta}\boldsymbol{\xi}_{m\beta} \ . \tag{1.27}$$

On reconnaît facilement dans la somme sur β le moment dipolaire de la molécule, $\mathbf{p}_m$. Du point de vue de la dérivation par rapport à l'argument de W, cette quantité est bien entendu une constante, et on peut écrire:

$$\rho_\ell^{(1)} = -\boldsymbol{\nabla}\cdot\sum_m \mathbf{p}_m W(\mathbf{r}-\mathbf{r}_m) = -\boldsymbol{\nabla}\cdot\mathbf{P} \ . \tag{1.28}$$

Comme nous le désirions, ce terme se calcule très simplement comme la divergence de la densité macroscopique de polarisation.

L'interprétation physique de ce terme est tout à fait transparente. Considérons un milieu, tel que celui représenté sur la figure 1.1, dont la densité de polarisation est uniforme, alignée avec Oz, dans une tranche d'espace d'épaisseur L perpendiculaire à Oz. On peut représenter chaque molécule comme un petit dipôle. On peut imaginer par exemple ces petits dipôles strictement alignés le long de la direction de polarisation. Si on prend un élément de volume quelconque dans le milieu, la densité moyenne de charges sera nulle parce que chaque tête positive d'une molécule quelconque est exactement compensée par la tête négative d'une molécule voisine. On trouve donc bien, avec cette image physique très simple, que la densité de charges dans le milieu uniformément polarisé est nulle. En revanche, si nous considérons un élément de volume recouvrant une interface du milieu avec l'extérieur, nous ne trouverons pas une charge nulle. Les charges à l'extrémité des molécules qui se trouvent immédiatement à la surface n'ont en effet aucune voisine pour les compenser. En un mot, nous trouverons des densités surfaciques de charge, positives sur la face d'entrée, négatives sur la face de sortie, ce qui est précisément le comportement de la divergence de $\mathbf{P}$ (au signe près) dans ce cas. Bien sûr, comme nous ne manipulons que des quantités moyennées à l'échelle macroscopique, nous avons une répartition continue de charges qui varie sur une échelle de l'ordre de s_0. Cette échelle étant négligeable par rapport aux dimensions caractéristiques du problème, on pourra toutefois assimiler cette distribution de charges à une répartition surfacique.

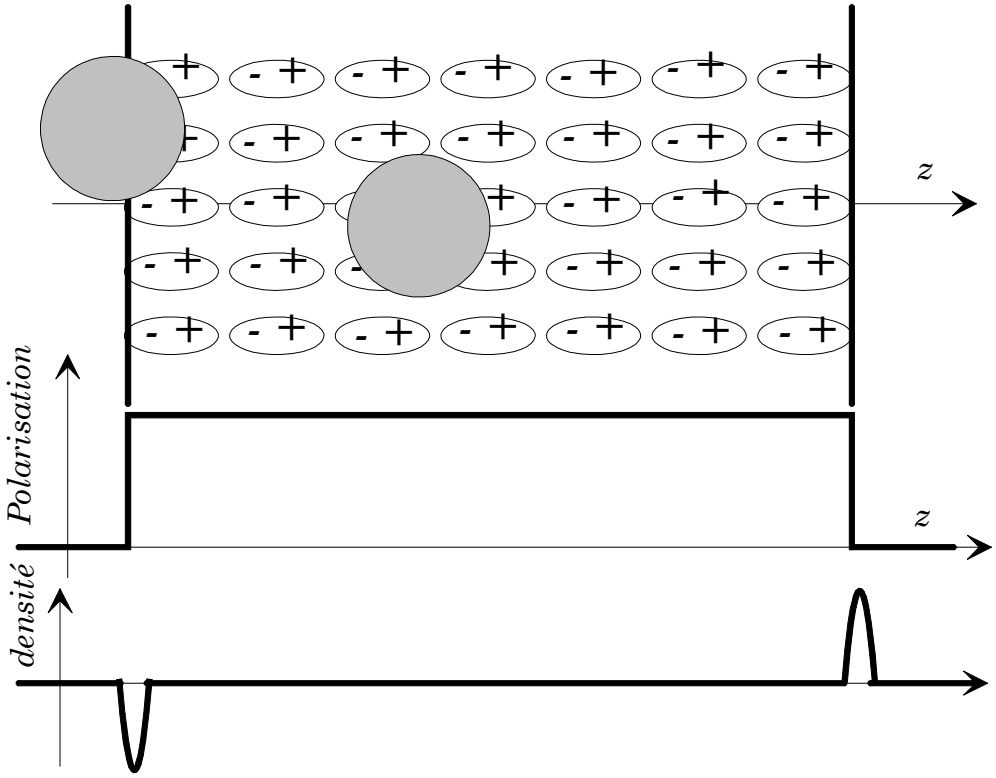

Figure 1.1: Répartition de charges dans un milieu uniformément polarisé. A l'intérieur du milieu, les charges de molécules voisines se compensent exactement. En revanche, sur les faces du milieu, apparaissent des densités surfaciques de charges. En grisé, on a représenté deux volumes de calcul de la densité moyenne de charges.

Ecrivons maintenant le terme d'ordre 2, dont il est d'ores et déjà évident qu'il fera intervenir la densité de quadripôles. Après quelques réarrangements, on a:

$$\rho_\ell^{(2)} = \frac{1}{2}\sum_j \partial_j \left\{ \sum_i \partial_i \left[\sum_{m,\beta} q_{m\beta}\xi_{m\beta i}\xi_{m\beta j} W(\mathbf{r}-\mathbf{r}_m) \right] \right\} \ . \tag{1.29}$$

Dans le crochet le plus intérieur, on reconnaît une composante de $\sum_m \mathbf{q}_m W$, c'est à dire de la densité moyenne de quadripôle $\boldsymbol{\mathcal{Q}}$. On a donc, en recourant un moment aux notations d'Einstein,

$$\rho_\ell^{(2)} = \frac{1}{2}\partial_j\partial_i \mathcal{Q}_{ij} \ . \tag{1.30}$$

En posant, comme nous l'avions déjà fait dans la quatrième partie,

$$\boldsymbol{\nabla}\cdot\boldsymbol{\mathcal{Q}}_j = \partial_i Q_{ji} \ , \tag{1.31}$$

on peut écrire:

$$\rho_\ell^{(2)} = \frac{1}{2}\boldsymbol{\nabla}\cdot(\boldsymbol{\nabla}\cdot\boldsymbol{\mathcal{Q}}) \ . \tag{1.32}$$

En regroupant enfin avec le terme d'ordre 1, nous allons trouver:

$$\rho_\ell = -\boldsymbol{\nabla}\cdot(\mathbf{P} - \frac{1}{2}\boldsymbol{\nabla}\cdot\boldsymbol{\mathcal{Q}}) \ , \tag{1.33}$$

le premier résultat essentiel de ce paragraphe.

Densité de courant

Nous allons maintenant procéder de la même manière pour la densité de courant lié. Nous développerons, dans l'expression (1.21), la fonction W en puissances de ξ. Toutefois, nous verrons qu'il suffit de pousser le développement à l'ordre un pour trouver les termes dipolaires magnétiques et quadripolaires électriques. Nous écrirons donc:

$$\mathbf{j}_\ell = \mathbf{j}_\ell^{(0)} + \mathbf{j}_\ell^{(1)} \ . \tag{1.34}$$

Le premier terme s'écrit simplement:

$$\mathbf{j}_\ell^{(0)} = \sum_{m,\beta} q_{m\beta}\dot{\boldsymbol{\xi}}_{m\beta}\, W(\mathbf{r}-\mathbf{r}_m) \ . \tag{1.35}$$

En isolant la somme sur β, dont la fonction W peut être sortie, on voit immédiatement intervenir:

$$\frac{d}{dt}\sum_\beta q_{m\beta}\boldsymbol{\xi}_{m\beta} = \frac{d\mathbf{p}_m}{dt} \ . \tag{1.36}$$

En constatant enfin que W ne dépend pas du temps, on a:

$$\mathbf{j}_\ell^{(0)} = \frac{\partial \mathbf{P}}{\partial t} \tag{1.37}$$

L'interprétation physique de ce terme est, là aussi, transparente. Si la polarisation du milieu varie, c'est que les molécules se déplacent ou se polarisent de façon différente. Cela résulte bien sûr en des déplacements de charges. Ils sont à l'échelle microscopiques, mais, comme toutes les molécules d'un même voisinage voient leur polarisation varier de la même manière, ces courants ont une moyenne non nulle à l'échelle macroscopique.

Le terme d'ordre 1 va nous demander un peu plus d'efforts. Il s'écrit en effet:

$$\mathbf{j}\ell^{(1)} = -\sum_{m,\beta} q_{m\beta} \left(\boldsymbol{\xi}_{m\beta} \cdot \boldsymbol{\nabla} W(\mathbf{r} - \mathbf{r}_m)\right) \dot{\boldsymbol{\xi}}_{m\beta} \ . \tag{1.38}$$

Comme nous l'avions fait en calculant le champ de dipôle magnétique et du quadripôle électrique, nous allons mettre artificiellement ce terme sous la forme d'un double produit vectoriel (l'analogie entre les deux calculs est bien sûr loin d'être fortuite). On écrit pour cela:

$$\begin{aligned}\left(\boldsymbol{\xi}_{m\beta} \cdot \boldsymbol{\nabla} W(\mathbf{r} - \mathbf{r}_m)\right) \dot{\boldsymbol{\xi}}_{m\beta} &= \frac{1}{2}\left[\left(\boldsymbol{\xi}_{m\beta} \cdot \boldsymbol{\nabla} W(\mathbf{r} - \mathbf{r}_m)\right) \dot{\boldsymbol{\xi}}_{m\beta} - \boldsymbol{\xi}_{m\beta}\left(\boldsymbol{\nabla} W(\mathbf{r} - \mathbf{r}_m) \cdot \dot{\boldsymbol{\xi}}_{m\beta}\right)\right] \\ &+ \frac{1}{2}\left[\left(\boldsymbol{\xi}_{m\beta} \cdot \boldsymbol{\nabla} W(\mathbf{r} - \mathbf{r}_m)\right) \dot{\boldsymbol{\xi}}_{m\beta} + \boldsymbol{\xi}_{m\beta}\left(\boldsymbol{\nabla} W(\mathbf{r} - \mathbf{r}_m) \cdot \dot{\boldsymbol{\xi}}_{m\beta}\right)\right]\end{aligned} \tag{1.39}$$

et donc:

$$\mathbf{j}\ell^{(1)} = \mathbf{j}_\ell^{(1)'} + \mathbf{j}_\ell^{(1)''} \ , \tag{1.40}$$

avec:

$$\mathbf{j}_\ell^{(1)'} = \frac{1}{2}\sum_{m,\beta} q_{m\beta} \boldsymbol{\nabla} W \times \left[\boldsymbol{\xi}_{m\beta} \times \dot{\boldsymbol{\xi}}_{m\beta}\right] \ , \tag{1.41}$$

l'autre terme étant défini pour compenser les termes additionnels introduits dans le premier. En isolant la somme sur m, en sortant la fonction W de la somme sur β, nous faisons apparaître le moment magnétique de la molécule m, $\frac{1}{2}\sum_\beta q_{m\beta} \boldsymbol{\xi}_{m\beta} \times \dot{\boldsymbol{\xi}}_{m\beta}$. Nous avons donc:

$$\mathbf{j}_\ell^{(1)'} = \sum_m \boldsymbol{\nabla} W \times \mathbf{m}_m \ . \tag{1.42}$$

Le moment magnétique de la molécule ne dépendant pas de la position dans l'espace, nous pouvons regrouper ces termes en faisant apparaître la densité macroscopique de moment magnétique:

$$\mathbf{j}_\ell^{(1)'} = \boldsymbol{\nabla} \times \boldsymbol{\mathcal{M}} \ . \tag{1.43}$$

L'interprétation physique de ce terme est très similaire à celle de la densité de charges. Considérons un barreau uniformément aimanté d'axe Oz, tel que celui représenté sur la figure 1.2. Nous pouvons remplacer chacun des dipôles magnétiques de ce barreau par une petite boucle carrée de courant, dont la taille est égale à la séparation entre boucles voisines. Toutes ces boucles sont parcourues par le même courant. Si nous considérons un point arbitraire dans le milieu (ou plutôt un voisinage d'extension s_0, les courants de boucles adjacentes s'annulent exactement: il n'y a pas de courants macroscopiques dans un milieu uniformément aimanté. En revanche, si nous considérons des points proches de la surface, les courants ne sont plus compensés. Il apparaît un courant de surface, ce qui est précisément l'aspect du rotationnel de l'aimantation comme on s'en convaincra aisément. Vu de l'extérieur, un barreau aimanté a donc la même répartition de courant qu'un solénoïde. On justifie ainsi l'évidente similitude des lignes de champ de ces deux objets. Notons que la densité de courant ainsi obtenue peut être considérable, de l'ordre de 10^6 A/m pour des aimants ordinaires. Elle résulte en effet de l'addition cohérente d'un très grand nombre de courants atomiques.

Traitons maintenant le second terme du courant à l'ordre 1. Sa composante j s'écrit:

$$\mathbf{j}_\ell^{(1)''}{}_j = -\frac{1}{2}\sum_{m,\beta} q_{m\beta}\left[\sum_i \xi_{m\beta i}\partial_i W \dot{\xi}_{m\beta j} + \sum_i \dot{\xi}_{m\beta i}\partial_i W \xi_{m\beta j}\right] \ . \tag{1.44}$$

On reconnaîtra aisément dans cette forme la composante j d'une dérivée par rapport au temps d'une quantité égale à:

$$-\frac{1}{2}\partial_i \sum_{m,\beta} q_{m\beta}\xi_{m\beta i}\xi_{m\beta j} \ , \tag{1.45}$$

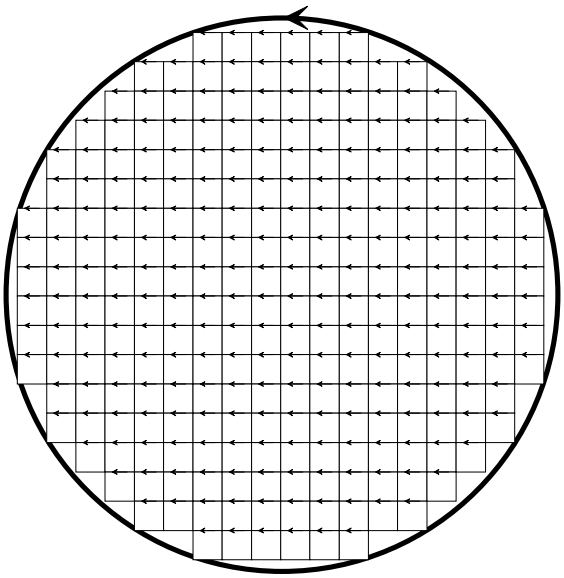

Figure 1.2: Section d'un barreau uniformément aimanté et représentation des courants microscopiques. Ils se compensent exactement, sauf au voisinage de la surface. Vu de l'extérieur, le barreau a le même champ magnétique qu'un solénoïde.

c'est à dire la dérivée par rapport au temps de la divergence du champ de quadripôle:

$$\mathbf{j}_\ell^{(1)''}{}_j = -\frac{\partial}{\partial t}\frac{1}{2}\boldsymbol{\nabla}\cdot\boldsymbol{\mathcal{Q}} \ . \tag{1.46}$$

Cette dérivée par rapport au temps, liée aux mouvements de charges imposés par une redistribution de la densité de quadripôles, se regroupe facilement avec le courant d'ordre 0 et on a finalement:

$$\mathbf{j}_\ell = \frac{\partial}{\partial t}\left(\mathbf{P} - \frac{1}{2}\boldsymbol{\nabla}\cdot\boldsymbol{\mathcal{Q}}\right) + \boldsymbol{\nabla}\times\boldsymbol{\mathcal{M}} \ . \tag{1.47}$$

Ceci est le deuxième résultat essentiel de ce paragraphe. Nous avons rempli totalement notre programme en exprimant ρ_ℓ et $\mathbf{j}_\ell$ en fonction des dérivées spatiales et temporelles de $\boldsymbol{\mathcal{M}}$ et de $\mathbf{P}-\frac{1}{2}\boldsymbol{\nabla}\cdot\boldsymbol{\mathcal{Q}}$. On constate d'ailleurs que les densités de dipôles et quadripôles électriques apparaissent toujours sous la forme de la même combinaison. Nous avons conservé la densité de quadripôles pour la cohérence de nos développements. Il est toutefois évident qu'elle jouera un rôle négligeable par rapport à la densité dipolaire, à moins que celle-ci ne soit nulle.

Nous avons totalement négligé, dans cette approche, les mouvements d'ensemble des molécules. Si on tenait compte correctement dans ce qui précède des termes en $\dot{\mathbf{r}}_m$, on montrerait que la densité de charges n'est pas modifiée($\dot{\mathbf{r}}_m$ n'y apparaît pas), mais qu'il apparaît un terme convectif dans la densité de courants liés:

$$\mathbf{j}_\ell = \frac{\partial}{\partial t}\left(\mathbf{P} - \frac{1}{2}\boldsymbol{\nabla}\cdot\boldsymbol{\mathcal{Q}}\right) + \boldsymbol{\nabla}\times\boldsymbol{\mathcal{M}} - \boldsymbol{\nabla}\times\left[\mathbf{V}\times(\mathbf{P} - \frac{1}{2}\boldsymbol{\nabla}\cdot\boldsymbol{\mathcal{Q}})\right] \ , \tag{1.48}$$

où $\mathbf{V}$ est la vitesse d'ensemble locale des molécules. On comprend bien en effet que le transport macroscopique d'une densité de polarisation soit équivalent à un courant.

1.3 Equations de Maxwell macroscopiques

Nous allons maintenant utiliser les expressions précédentes des charges et courants liés pour réécrire les équations de Maxwell à l'échelle macroscopique.

1.3.1 Déplacement électrique, Induction magnétique

Les équations de Maxwell homogènes restent bien sûr inchangées dans les opérations de moyennage:

$$\boldsymbol{\nabla} \times \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t} \tag{1.49}$$

$$\boldsymbol{\nabla} \cdot \mathbf{B} = 0 \ . \tag{1.50}$$

L'équation de Maxwell–Gauss se réécrit:

$$\boldsymbol{\nabla} \cdot \mathbf{E} = (\rho + \rho_\ell)/\epsilon_0 \ , \tag{1.51}$$

où ρ est, rappelons–le, la densité de charges libres moyennée. En remplaçant ρ_ℓ par l'expression (1.33), on voit que cette équation peut se mettre sous la forme:

$$\boldsymbol{\nabla} \cdot \mathbf{D} = \rho \ , \tag{1.52}$$

où le vecteur $\mathbf{D}$, appelé déplacement électrique, est défini par:

$$\mathbf{D} = \epsilon_0 \mathbf{E} + \mathbf{P} - \frac{1}{2} \boldsymbol{\nabla} \cdot \boldsymbol{\mathcal{Q}} \ . \tag{1.53}$$

Bien sûr l'introduction de ce vecteur ne fait, à ce stade, que simplifier les écritures. Il faudrait pourvoir écrire explicitement les densités de polarisation pour lui donner un sens utile.

Finalement, l'équation de Maxwell ampère peut s'écrire, en remplaçant les courants liés par leur expression (1.47),

$$\boldsymbol{\nabla} \times \mathbf{B} = \mu_0 \mathbf{j} + \frac{1}{\epsilon_0 c^2} \left[\frac{\partial}{\partial t} \left(\mathbf{P} - \frac{1}{2} \boldsymbol{\nabla} \cdot \boldsymbol{\mathcal{Q}} \right) + \boldsymbol{\nabla} \times \boldsymbol{\mathcal{M}} \right] + \frac{1}{c^2} \frac{\partial \mathbf{E}}{\partial t} \ . \tag{1.54}$$

En regroupant les dérivées par rapport au temps, on peut faire apparaître le déplacement électrique $\mathbf{D}$. En regroupant les termes en rotationnel, on peut finalement mettre cette équation sous la forme:

$$\boldsymbol{\nabla} \times \mathbf{H} = \mathbf{j} + \frac{\partial \mathbf{D}}{\partial t} \ , \tag{1.55}$$

en introduisant l'induction magnétique $\mathbf{H}$ définie par:

$$\mathbf{H} = \frac{\mathbf{B}}{\mu_0} - \boldsymbol{\mathcal{M}} \ . \tag{1.56}$$

Nous avons finalement obtenu, par de simples jeux de réécriture, quatre équations qui ont à peu près la forme des équations de Maxwell et qui ne font apparaître que les densités de charges et de courants libres, qui sont sous le contrôle de l'expérimentateur. La difficulté est bien sûr que ces équations sont écrites en fonction de quatre champs: les champs électriques et magnétiques traditionnels, le déplacement électrique et l'induction magnétique (ces noms ont été donné historiquement très tôt dans le développement de l'électromagnétisme et ont été conservés depuis). Il n'y a donc pas de solution unique à cet ensemble d'équations, sous déterminées. Ce n'est qu'en précisant les relations entre les densités de polarisation et les champs que nous pourrons exprimer les deux nouveaux champs en fonction des champs électriques et magnétiques et obtenir un ensemble de quatre équations ne portant que sur deux champs de vecteurs, qu'il devrait être possible de résoudre ensuite.

1.3.2 Considérations énergétiques

Nous terminerons ce chapitre par quelques brèves remarques sur l'énergétique. Il est clair que des densités d'énergie doivent être associées aux densités de polarisation et que le transport de polarisation doit contribuer au transport d'énergie. En fait, les bilans d'énergie dans la matière sont complexes puisqu'ils doivent faire intervenir tout autant les énergies associés aux champs et polarisations que l'énergie de la matière elle même. Des phénomènes comme l'électrostriction, par exemple, (la matière devient plus dense quand on lui applique un champ magnétique) ne peuvent être correctement compris que dans une approche thermodynamique globale.

Nous ne tenterons pas cette approche ici et nous nous contenterons d'établir, à partir des équations de Maxwell macroscopiques, une équation ayant une forme similaire au bilan d'énergie dans le vide (vecteur de Poynting). Nous postulerons alors l'identification des différents termes de cette équation à une densité d'énergie électromagnétique et à un vecteur décrivant le flux d'énergie.

Nous procéderons en fait comme on le fait pour les équations de Maxwell dans le vide. A partir de Maxwell Ampère, on écrit:

$$\mathbf{j} = \boldsymbol{\nabla} \times \mathbf{H} - \frac{\partial \mathbf{D}}{\partial t} \tag{1.57}$$

et on en déduit la densité de puissance cédée par les champs aux courants libres, $\mathbf{j} \cdot \mathbf{E}$ (l'expression de la force de Lorentz s'exerçant sur les porteurs libres fait bien sûr intervenir les champs ordinaires). On a

$$\mathbf{j} \cdot \mathbf{E} = \mathbf{E} \cdot (\boldsymbol{\nabla} \times \mathbf{H}) - \mathbf{E} \cdot \frac{\partial \mathbf{D}}{\partial t} \ . \tag{1.58}$$

En remarquant que:

$$\mathbf{E} \cdot (\boldsymbol{\nabla} \times \mathbf{H}) = \mathbf{H} \cdot (\boldsymbol{\nabla} \times \mathbf{E}) - \boldsymbol{\nabla} \cdot (\mathbf{E} \times \mathbf{H}) \ , \tag{1.59}$$

et en utilisant la première équation de Maxwell, on a:

$$\mathbf{j} \cdot \mathbf{E} + \mathbf{E} \cdot \frac{\partial \mathbf{D}}{\partial t} + \mathbf{H} \cdot \frac{\partial \mathbf{B}}{\partial t} + \boldsymbol{\nabla} \cdot (\mathbf{E} \times \mathbf{H}) = 0 \ . \tag{1.60}$$

Cette équation a exactement la forme de l'équation bilan dans le vide. $\mathbf{j} \cdot \mathbf{E}$ étant la puissance cédée par le champ aux charge libres et donc aussi la puissance cédée par les générateurs au champ, on peut interpréter

$$\boldsymbol{\Pi} = \mathbf{E} \times \mathbf{H} \tag{1.61}$$

comme le vecteur dont le flux à travers une surface représente le transport d'énergie à travers cette surface. Il s'agit donc du vecteur de Poynting dans la matière. Finalement, le terme

$$\mathbf{E} \cdot \frac{\partial \mathbf{D}}{\partial t} + \mathbf{H} \cdot \frac{\partial \mathbf{B}}{\partial t} \tag{1.62}$$

pourrait être interprété comme la dérivée par rapport au temps de la densité d'énergie électromagnétique. Dans le cas général, sans connaître le lien entre polarisation et champ, on ne peut préciser davantage la densité d'énergie électromagnétique. On peut cependant analyser un peu plus en détail la signification physique des termes, en particulier électriques. Le premier terme peut en effet s'écrire

$$\mathbf{E} \cdot \frac{\partial \mathbf{D}}{\partial t} = \frac{\partial}{\partial t} \epsilon_0 \frac{E^2}{2} + \mathbf{E} \cdot \frac{\partial \mathbf{P}}{\partial t} \ . \tag{1.63}$$

La première dérivée n'est autre que celle de la densité d'énergie électrostatique ordinaire. Le deuxième terme peut s'écrire:

$$\mathbf{E} \cdot \frac{\partial \mathbf{P}}{\partial t} = \mathbf{E} \cdot \mathbf{j}_\ell \ . \tag{1.64}$$

si on suppose le matériau dépourvu de propriétés magnétiques et donc le courant lié seulement à la variation temporelle de la polarisation électrique. Il s'agit donc simplement de la puissance

que le champ fournit à la matière pour changer sa polarisation. Ce changement de polarisation s'accompagne effectivement de courants macroscopiques produits par des forces de Lorentz qui travaillent. L'interprétation physique de la partie électrique de cette équation bilan est donc très claire: la variation dans le temps de l'énergie électrique compense la puissance cédée aux courants macroscopiques ou aux courants liés résultant des changements de polarisation.

Chapitre 2

Réponse linéaire

Nous allons pouvoir donner un sens physique aux équations établies dans le chapitre précédent en précisant les relations entre les densités de polarisation et les champs. Il paraît assez naturel que, dans la plupart des milieux, les polarisations électriques ou magnétiques répondent aux champs appliqués. De façon très générale, on doit pouvoir, avec un modèle microscopique du milieu, écrire les densités de polarisation en fonction des champs. Si les champs appliqués ne sont pas trop grand, cette relation est a priori linéaire. Dans la plupart des matériaux, à l'exception notable des ferromagnétiques et des ferroélectriques, les polarisations s'annulent en l'absence de champ. On doit donc pouvoir écrire une relation linéaire homogène entre les polarisations et les champs. Cette relation doit se mettre sous la forme d'un produit de convolution faisant intervenir la fonction de Green du système, ou encore sa réponse percussionnelle. Nous retrouverons donc des points du formalisme assez semblables à ce que nous avions utilisé pour établir l'expression des potentiels retardés dans la troisième partie de ce cours. Plutôt qu'un produit de convolution, nous préférerons écrire de simples relations de proportionnalité en introduisant les composantes de Fourier des différentes quantités. Nous définirons alors la notion de susceptibilité. L'ensemble de ces notions fera la matière du premier paragraphe.

Dans le second paragraphe, nous passerons en revue rapidement quelques modèles microscopiques de polarisabilité électrique ou magnétique. C'est en effet par un mécanisme de polarisation, comme celui que nous avons déjà rencontré pour le modèle de Thomson, que chaque molécule acquiert un dipôle sous l'influence du champ appliqué.

Si le milieu était très dilué, il n'y aurait aucun problème pour exprimer la susceptibilité, quantité macroscopique, en fonction de la polarisabilité, quantité microscopique. En fait, les choses ne sont pas aussi simples. Le champ "vu" par chaque molécule est la somme du champ extérieur, macroscopique, et du champ rayonné par les molécules immédiatement voisines. Ce champ étant un champ microscopique, il n'est pas convenablement décrit dans notre formalisme. Nous apprendrons, dans le troisième paragraphe de ce chapitre, comment on peut contourner cette difficulté et exprimer, au moyen d'hypothèses très générales et très bien vérifiées, les champs "locaux" en fonction des champs macroscopiques. Nous pourrons alors écrire simplement les susceptibilités en fonction des polarisabilités et de la densité numérique du milieu.

Avec ces nouveaux outils, nous reviendrons sur les bilans énergétiques dans la matière. Nous montrerons en particulier que la partie imaginaire de la susceptibilité décrit très simplement les échanges d'énergie entre champ et matière. Nous verrons que, dans la plupart des matériaux, la matière absorbe le champ. Nous montrerons que, à certaines conditions, un milieu atomique renforce le champ incident. Nous retrouverons ainsi, de manière très élémentaire, le principe de l'émission stimulée, à la base du fonctionnement des lasers.

Le dernier paragraphe de ce chapitre sera consacré à l'établissement des relations de Kramers–Krönig. Nous verrons que la simple causalité (la polarisation ne peut répondre avant l'application du champ) impose des relations très fortes entre les parties réelles et imaginaires de la susceptibilité, entre dispersion et absorption. Ces relations sont très générales dans la théorie de la réponse linéaire ou des

fonctions de Green et peuvent être appliquées à bien d'autres domaines que l'électromagnétisme dans la matière.

2.1 Susceptibilités

Nous essaierons d'abord d'écrire les relations linéaires les plus générales possibles entre polarisations et champ. Nous commencerons par formuler quelques hypothèses raisonnables. Elles ne sont pas toutes indispensables, mais elles facilitent beaucoup l'algèbre sans beaucoup restreindre la généralité. Nous supposerons donc que:

- $\mathbf{P}$ et $\mathbf{j}$ ne dépendent que de $\mathbf{E}$. Pour la première hypothèse, nous négligerons les effets magnétiques sur la polarisabilité électrique des molécules. Les vitesses des électrons dans les molécules, principalement responsables de la polarisabilité électrique, sont en effet faibles par rapport à la vitesse de la lumière. L'action des forces magnétiques est donc bien négligeable. Pour la seconde hypothèse, nous l'introduisons pour décrire la conductivité du matériau. Nous nous pencherons surtout sur la propagation d'ondes dans les milieux homogènes, d'où les générateurs sont exclus, et où les charges ne se mettent en mouvement que sous l'action du champ électrique. Là aussi, nous négligerons l'influence des forces magnétiques et des phénomènes tels que la magnétorésistance ou l'effet Hall. Ils ne sont guère observés que dans les semi–conducteurs, et pourront être traités en perturbation par rapport à ce que nous établirons ici.

- $\boldsymbol{\mathcal{M}}$ ne dépend que de $\mathbf{B}$. Cette hypothèse est symétrique de la précédente.

- les densités de quadripôles sont tout à fait négligeables. Tout milieu présentant au moins une polarisabilité électrique dipolaire induite, cette hypothèse est plus que raisonnable.

- toutes les polarisations s'annulent avec le champ qui en est la cause. On néglige donc les matériaux où apparaissent des polarisations permanentes, ferromagnétiques ou ferroélectriques. Ces matériaux présentant tous de l'hystérésis (l'état dépend de toute l'histoire du matériau), ils ne peuvent être abordés dans un cadre simple.

- la relation entre polarisations et champs est locale. La polarisation ne dépend que du champ, pris à un instant arbitraire, au même endroit. Nous négligerons phénomènes de transport et corrélations diverses.

- la relation est causale. La polarisation ne dépend du champ que dans le passé.

Les relations que nous allons écrire ayant les mêmes formes pour $\mathbf{P}, \mathbf{j}$ et $\boldsymbol{\mathcal{M}}$, nous ne les écrirons systématiquement que pour $\mathbf{P}$. Nous voulons une relation linéaire homogène entre $\mathbf{P}$ et $\mathbf{E}$. Elle peut s'écrire en termes d'une fonction de Green. $\mathbf{G}(\mathbf{r}, t)$ étant, à des facteurs dimensionnels près, la polarisation créée au point $\mathbf{r}$ par un champ impulsionnel en $t = 0$, on pourra écrire:

$$\mathbf{P}(\mathbf{r}, t) = \frac{\epsilon_0}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \mathbf{G}(\mathbf{r}, \tau) \cdot \mathbf{E}(t - \tau)\, d\tau \ . \tag{2.1}$$

Il n'y a aucune raison a priori pour que $\mathbf{P}$ et $\mathbf{E}$ aient la même direction. Un matériau cristallin, par exemple, ne répond pas de manière identique dans les différentes directions propres de sa maille cristalline. La fonction de Green $\mathbf{G}$ est donc a priori un tenseur de rang 2, qui ne se réduira à un scalaire que pour des milieux isotropes. La causalité, pour sa part, implique que $\mathbf{G}$ s'annule pour $\tau < 0$. On vérifiera d'autre part que, avec les normalisations employées, $\mathbf{G}$ a la dimension d'une fréquence (son intégrale sur le temps est sans dimension).

Cette relation de convolution est d'un maniement difficile. On peut la transformer en une simple relation de proportionnalité en introduisant les transformées de Fourier temporelles. En posant

$$\mathbf{P}(\mathbf{r},t) = \frac{1}{\sqrt{2\pi}} \int \mathbf{P}_\omega(\mathbf{r}) e^{-i\omega t}\, d\omega \tag{2.2}$$

et une définition équivalente pour $\mathbf{E}$, on peut écrire:

$$\begin{aligned} \mathbf{P}_\omega(\mathbf{r}) &= \frac{1}{\sqrt{2\pi}} \int \mathbf{P}(\mathbf{r},t) e^{i\omega t}\, dt \\ &= \frac{\epsilon_0}{2\pi} \int G(\mathbf{r},\tau) \cdot \mathbf{E}(\mathbf{r},t-\tau) e^{i\omega t}\, dt d\tau \\ &= \frac{\epsilon_0}{(2\pi)^{3/2}} \int G(\mathbf{r},\tau) \cdot \mathbf{E}_\omega(\mathbf{r}) e^{i(\omega-\omega')t} e^{i\omega'\tau}\, dt d\tau d\omega' \end{aligned} \tag{2.3}$$

En remarquant alors que:

$$\int dt e^{i(\omega-\omega')t} = 2\pi\delta(\omega-\omega')\ , \tag{2.4}$$

on a

$$\mathbf{P}_\omega(\mathbf{r}) = \epsilon_0 \boldsymbol{\chi}^e(\omega,\mathbf{r}) \cdot \mathbf{E}_\omega\ , \tag{2.5}$$

où $\boldsymbol{\chi}^e$ est évidemment la transformée de Fourier de la fonction de Green:

$$\boldsymbol{\chi}^e(\omega,\mathbf{r}) = \frac{1}{\sqrt{2\pi}} \int \mathbf{G}(\tau,\mathbf{r}) e^{i\omega t}\, dt\ . \tag{2.6}$$

Nous appellerons $\boldsymbol{\chi}^e$ la susceptibilité du matériau. L'exposant e a été ajouté pour distinguer plus tard la susceptibilité électrique de la susceptibilité magnétique. C'est une fonction de la fréquence, qui exprime la relation de proportionnalité entre les composantes de fréquence de $\mathbf{P}$ et de $\mathbf{E}$. C'est aussi une fonction de la position, si le matériau n'est pas homogène. C'est enfin, comme G, une quantité tensorielle de rang 2, sans dimensions. On notera à ce propos que $\mathbf{P}_\omega$ et $\mathbf{E}_\omega$ ont la dimension d'une polarisation ou d'un champ électrique multipliés par un temps.

Nous considérerons surtout des milieux isotropes. Nous écrirons donc à partir de maintenant $\boldsymbol{\chi}^e$ comme un simple scalaire χ^e. La plupart de nos raisonnements pourraient être généralisés au cas tensoriel par une simple réécriture. χ^e est a priori une quantité complexe. La polarisation ne répondant pas instantanément au champ, il doit exister un déphasage non nul entre leurs composantes de Fourier. Nous aurons souvent à distinguer les parties réelles et imaginaires et écrirons donc χ^e sous la forme:

$$\chi^e = \chi^{e'} + i\chi^{e''}\ , \tag{2.7}$$

où $\chi^{e'}$ et $\chi^{e''}$ sont des quantités réelles.

Ce que nous venons de faire pour $\mathbf{P}$, nous pouvons aussi le faire pour $\boldsymbol{\mathcal{M}}$ et introduire une susceptibilité magnétique χ^m définie par:

$$\boldsymbol{\mathcal{M}}_\omega = \frac{\chi^m(\mathbf{r},\omega)}{\mu_0} \mathbf{B}_\omega\ . \tag{2.8}$$

avec ces notations, la susceptibilité magnétique est elle aussi sans dimensions. Enfin, nous pouvons introduire une conductivité à la fréquence ω en posant:

$$\mathbf{j}_\omega = \sigma_\omega(\mathbf{r}) \mathbf{E}_\omega\ . \tag{2.9}$$

Notons qu'à la différence des précédentes, cette quantité n'est pas sans dimensions. Pour des raisons principalement historiques, on écrit une conductivité plutôt qu'un susceptibilité à proprement parler. Notons que cette définition peut être associée à l'équation de conservation de la charge écrite en termes des composantes de Fourier:

$$i\omega\rho_\omega = \boldsymbol{\nabla} \cdot \mathbf{j}_\omega\ . \tag{2.10}$$

Nous pouvons ainsi déterminer aussi la densité de charges libres en fonction du champ électrique.

2.2 Polarisabilité électrique

De manière évidente, si ces relations de susceptibilités existent, c'est que chaque molécule du milieu prend un petit moment dipolaire magnétique sous l'influence du champ magnétique et un petit moment dipolaire électrique sous l'influence du champ électrique. On doit donc pouvoir déduire les susceptibilités d'un modèle microscopique. Nous allons nous employer, dans ce paragraphe et le suivant, à passer rapidement en revue des modèles microscopiques de polarisabilité. Nous attendrons le dernier chapitre de cette partie pour donner un modèle de conductivité, plus phénoménologique que physique. Nous commencerons donc par la polarisabilité électrique. Nous chercherons donc à écrire le dipôle moyen pris par une molécule $\mathbf{p}_m$ sous l'influence d'un champ électrique $\mathbf{E}_m$ sous la forme:

$$\mathbf{p}_m = \epsilon_0 \alpha_m^e \mathbf{E}_m \ . \tag{2.11}$$

Nous noterons le champ $\mathbf{E}_m$. Il s'agit en effet du champ microscopique vu par la molécule, non du champ macroscopique qui intervient dans les équations de Maxwell moyennées. Notons aussi qu'il s'agit d'une réponse moyenne. Dans les susceptibilités n'apparaissent que des quantités macroscopiques, moyennées sur un grand nombre de molécules. Nous pourrons par exemple recourir à la thermodynamique pour écrire des moyennes d'ensemble sur un très grand nombre de molécules sous l'action conjointe du champ appliqué et de l'agitation thermique. La polarisabilité peut se présenter sous deux formes: polarisabilité induite ou d'orientation.

2.2.1 Polarisabilité induite

Les atomes ou les molécules plongés dans un champ électrique oscillant acquièrent un moment dipolaire induit. Le modèle le plus simple est donné par le modèle de Thomson de l'électron élastiquement lié, dont nous avons établi la surprenante validité dans le domaine quantique. On se reportera donc à la quatrième partie pour l'expression de la polarisabilité α_i associée.

2.2.2 Polarisablilité d'orientation

En plus de la polarisabilité induite, certaines molécules présentent un moment dipolaire permanent (interdit en revanche pour les atomes). C'est par exemple le cas pour les molécules des solvants polaires comme l'eau, l'ammoniaque ou HCl. En l'absence de champ appliqué, les orientations des dipôles sont aléatoires. En moyenne, il n'existe pas de polarisation [1]. Sous l'influence d'un champ électrique et de la relaxation, les dipôles du milieu vont tendre à s'aligner avec le champ. Il apparaîtra donc un dipôle induit, proportionnel au champ si celui-ci n'est pas trop rapide.

Il est évident que, pour un champ harmonique, le dipôle moléculaire moyen ne pourra pas suivre le champ électrique s'il oscille à trop haute fréquence. Il faut que la période du champ soit plus longue que le temps de relaxation caractéristique du milieu, c'est à dire le temps pour qu'un dipôle, en moyenne revienne à l'équilibre thermodynamique. Ce temps de relaxation est en général très court, de l'ordre du temps entre collisions dans les liquides, donc inférieur à la picoseconde. La polarisation a donc le temps de s'établir dès que la champ a une longueur d'onde supérieure à celles de l'infrarouge proche. On peut alors se contenter d'un modèle statique, que nous allons détailler plus loin. A très haute fréquence, la molécule ne répond pratiquement plus (en moyenne statistique) au champ appliqué et la polarisabilité s'annule. Entre les deux régimes, c'est à dire dans le visible et proche infrarouge, on a une situation plus complexe où le module et la phase de la polarisabilité (qui est évidemment une quantité complexe) évoluent rapidement.

Nous allons maintenant calculer la polarisabilité statique (ou basse fréquence) d'un ensemble de molécules polaires. Nous traiterons le problème sans recourir à la mécanique quantique, ce qui serait

[1] Dans les matériaux ferroélectriques, les dipôles sont associés à une déformation permanente de la maille cristalline, et il peut apparaître une polarisation en l'absence de champ appliqué. Nous avons explicitement exclu ce type de matériau de notre étude

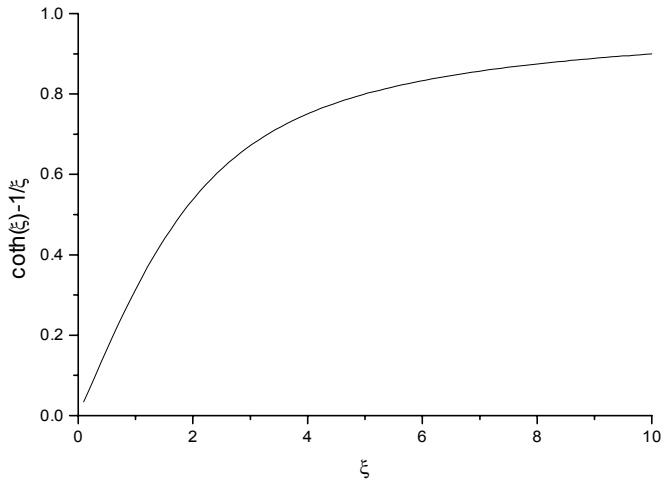

Figure 2.1: Dipôle moyen en fonction du paramètre xi.

indispensable pour une approche rigoureuse. Le dipôle électrique, en projection sur un axe, est une quantité quantifiée. Nous verrons, pour la polarisation d'orientation magnétique, comment tenir compte en principe de cette quantification. Elle ne modifie pas l'aspect qualitatif des choses et n'altère qu'un facteur numérique proche de l'unité portant sur la polarisabilité. Pour une discussion qualitative, nous pourrons donc nous contenter d'une approche classique.

Il est évident par symétrie que le dipôle est aligné avec le champ, que nous prendrons selon Oz. Si p_0 est le module du dipôle moléculaire permanent, l'énergie d'interaction d'une molécule dont l'axe (et donc le dipôle) font un angle θ avec l'axe Oz est simplement $H = -p_0 E_m \cos\theta$. On peut alors utiliser la statistique de Boltzmann pour calculer la composante selon z du dipôle moyen:

$$\overline{p} = p_0 \int \frac{1}{Z} e^{-\beta H} \cos\theta \ , \tag{2.12}$$

avec $\beta = 1/kT$. Z est la fonction de partition. L'intégrale porte sur toutes les configurations possibles, c'est à dire sur tous les angles θ, ϕ des coordonnées sphériques. L'intégration sur ϕ est un simple facteur 2π, absorbé par un facteur identique dans la fonction de partition. On a donc:

$$\overline{p} = p_0 \frac{\int \sin\theta \, d\theta \ \cos\theta \exp(p_0 E_m \cos\theta / kT)}{\int \sin\theta \, d\theta \ \exp(p_0 E \cos\theta / kT)} \ . \tag{2.13}$$

Ces intégrales se calculent de façon triviale. En posant

$$\xi = \frac{p_0 E_m}{kT} \ , \tag{2.14}$$

rapport de l'énergie du dipôle dans le champ à l'énergie thermique moyenne, on trouve:

$$\overline{p} = p_0 \left(\coth\xi - \frac{1}{\xi} \right) \ . \tag{2.15}$$

La fonction $\coth\xi - \frac{1}{\xi}$ est représentée sur la figure 2.1 La relation que nous obtenons ainsi est non linéaire. Pour des très grands champs, tels que l'énergie du dipôle dans le champ soit beaucoup plus grande que l'énergie thermique, le dipôle moyen sature à une valeur p_0. En fait toutes les polarisabilités d'orientation présentent cette saturation triviale quand toutes les molécules du milieu sont alignées.

Nous ne trouverons de polarisabilité linéaire qu'en allant dans le domaine des tout petits champs électriques. Si $\xi \ll 1$, nous pouvons en effet faire un développement limité en puissances de ξ. Après quelques manipulations algébriques élémentaires, on trouvera:

$$\overline{p} = p_0 \frac{\xi}{3} \ , \tag{2.16}$$

et donc une polarisabilité d'orientation:

$$\alpha_o = \frac{p_0^2}{3\epsilon_0 kT} \ , \tag{2.17}$$

homogène à un volume comme il se doit. L'hypothèse de champ faible n'est pas trop contraignante. Si nous prenons un dipôle de l'ordre de 3 debyes, 10^{-29} Cm, dipôle très grand par rapport aux dipôles moléculaires ordinaires, la limite des champs faibles correspond à $E_m \ll 10^9$ V/m. Pour de tels champs, la plupart des matériaux ont déjà donné lieu à un claquage (les rigidités diélectriques typiques des matériaux les plus résistants, comme le Téflon, sont de quelques dizaines de kV/mm, ou 10^7 V/m)!

Comparons, en ordres de grandeur, la polarisabilité d'orientation à la polarisabilité induite du modèle de Thomson. En posant l'ordre de grandeur de $p_0 = a_0 e$ où e est la charge élémentaire et a_0 le rayon de Bohr, on pourra écrire:

$$\alpha_0 \simeq \frac{a_0^2 e^2}{3\epsilon_0 kT} = \frac{e^2}{8\pi\epsilon_0 a_0} \frac{1}{kT} \frac{8\pi}{3} a_0^3 \ . \tag{2.18}$$

Dans la première fraction du membre de droite, on reconnaît l'énergie de liaison atomique, de l'ordre de la constante de Rydberg R. Le dernier terme est de l'ordre du volume de la molécule, qui est précisément l'ordre de grandeur de la polarisabilité induite α_i dans la limite statique, qui s'applique largement ici. On a donc finalement:

$$\frac{\alpha_o}{\alpha_i} = \frac{R}{kT} \simeq 10^3 \ . \tag{2.19}$$

Pour toute température réaliste, la polarisabilité d'orientation domine largement la polarisabilité induite. Celle-ci ne jouera donc de rôle que pour les molécules non polaires, ou à très haute fréquence. Dans ce cas la polarisabilité d'orientation est pratiquement nulle. La polarisabilité induite, qui met en jeu des constantes de temps atomiques et non des constantes de temps de thermalisation, peut encore être importante.

2.3 Polarisabilités magnétiques

Nous chercherons ici à écrire le moment magnétique moléculaire moyen en fonction du champ appliqué, sous la forme:

$$\mathbf{m} = \frac{\alpha^m}{\mu_0} \mathbf{B}_m \ . \tag{2.20}$$

Avec ces notations, α^m a bien la dimension d'un volume. Là aussi $\mathbf{B}_m$ est le champ "vu" par la molécule m qu'il nous faudra déterminer en fonction des champs macroscopiques.

2.3.1 Diamagnétisme

Modèle classique

La première source de polarisabilité magnétique est une polarisabilité induite que présentent toutes les molécules ou atomes n'ayant pas de dipôle magnétique permanent. Pour comprendre l'origine de cette polarisabilité, nous considérerons un modèle atomique très simple. Un atome est constitué

de deux électrons orbitant autour du noyau sur une orbite de rayon r, perpendiculaire à l'axe Oz, parcourue en sens inverse par les deux électrons. Si nous ne considérions qu'un électron (atome d'hydrogène), nous aurions un moment magnétique orbital permanent. Avec ce modèle très naïf (il néglige l'interaction coulombienne entre les deux électrons) de l'atome d'hélium, on n'a pas de moment magnétique permanent parce que les moments associés aux deux électrons se compensent exactement.

Appliquons à cet atome un champ magnétique uniforme, statique, $\mathbf{B}_m$ dirigé selon Oz. Ce champ doit croître, lentement à l'échelle des temps atomiques, depuis la valeur nulle. Comme tout champ magnétique variable, il génère un champ électrique. Celui-ci est visiblement, en coordonnées cylindriques, dirigé selon $\mathbf{u}_\theta$: il tourne autour de l'axe. Ce champ électrique va donc ralentir un des électrons et accélérer l'autre. Il va, en un mot, briser la compensation des moments magnétiques orbitaux et faire apparaître un moment magnétique induit [2].

Précisons cet argument. Le potentiel vecteur $\mathbf{A}$ s'écrit, toujours en coordonnées cylindriques:

$$\mathbf{A} = \frac{1}{2}B(t)r\mathbf{u}_\theta \ , \tag{2.21}$$

où $B(t)$ est le module instantané du champ, passant adiabatiquement de 0 à B_m. Le champ électromoteur s'écrit donc:

$$\mathbf{E} = -\frac{1}{2}r\frac{dB}{dt}\mathbf{u}_\theta \ . \tag{2.22}$$

En supposant, ce qui est une excellente approximation, que l'orbite des électrons reste invariante et que seul le module de la vitesse de l'électron change, on trouve que cette modification de vitesse est (pour l'électron tournant dans le sens direct autour de B):

$$\Delta v = -\frac{q}{2m}r\int \frac{dB}{dt} \ , \tag{2.23}$$

l'intégrale portant sur toute la phase de variation du champ magnétique et q étant la charge (négative) de l'électron. On a donc:

$$\Delta v = -\frac{qB_m}{2m}r \ , \tag{2.24}$$

indépendamment de la loi de variation de B. Il s'agit donc d'un résultat universel. L'autre électron a évidemment une variation de vitesse opposée. Il en résulte une variation des moments magnétiques associés aux deux électrons et un moment magnétique global:

$$m = -\frac{q^2r^2}{4m}B_m \ . \tag{2.25}$$

On en déduit la polarisabilité diamagnétique:

$$\alpha_d^m = -\frac{q^2\mu_0}{4m}\sum_i r_i^2 \ , \tag{2.26}$$

où la somme est à étendre à tous les électrons (au moins deux) de l'atome.

Nous avons bien une polarisabilité linéaire. Elle a été calculée ici pour un champ statique. Les constantes de temps de mise à l'équilibre des orbitales atomiques étant dans le domaine optique, cette expression de la polarisabilité diamagnétique doit rester valable jusqu'à des fréquences très élevées. La polarisabilité diamagnétique est négative. Le moment induit est opposé au champ magnétique. Cela implique qu'une telle molécule aura, pour minimiser l'énergie, tendance à fuir les régions de champs forts. C'est effectivement le comportement des substances diamagnétiques, dominées par ce type de polarisabilité.

[2] La variation de vitesse des électrons lors du branchement du champ magnétique est souvent appelée "effet bétatron" en référence à l'un des premiers accélérateurs de particules. On utilisait un champ magnétique croissant pour accélérer des électrons orbitant dans ce champ. L'énergie finale étant limitée par le champ maximal, les performances étaient plus que limitées et le principe fut rapidement abandonné

Modèle quantique

Le modèle très naïf du paragraphe précédent ne peut nous satisfaire. Pour essayer de préciser les choses, nous allons nous pencher sur un atome à un électron plongé dans un champ magnétique. Nous essaierons de calculer le dipôle magnétique induit.

Le hamiltonien de l'atome dans le champ est, comme nous l'avons déjà vu dans la partie précédente:

$$H = \frac{(\mathbf{P} - q\mathbf{A})^2}{2m} + V(\mathbf{R}) \ , \tag{2.27}$$

où $\mathbf{P}$ et $\mathbf{R}$ sont les opérateurs position et impulsion, et $\mathbf{A}$ le potentiel vecteur, dépendant de $\mathbf{R}$. V est simplement le potentiel coulombien. Le moment magnétique peut s'écrire:

$$\mathbf{m} = \frac{q}{2}\mathbf{R} \times \mathbf{v} \ , \tag{2.28}$$

où $\mathbf{v}$ est l'opérateur vitesse, qui s'écrit, comme en mécanique analytique classique:

$$\mathbf{v} = \frac{\mathbf{P} - q\mathbf{A}}{m} \ . \tag{2.29}$$

En notant que $\mathbf{R} \times \mathbf{P} = \mathbf{L}$ est le moment cinétique, on a donc:

$$\mathbf{m} = \frac{q}{2m}\left[\mathbf{L} - q\mathbf{R} \times \mathbf{A}\right] \ . \tag{2.30}$$

Evaluons la valeur moyenne de cet opérateur dans l'état fondamental $|1S\rangle$ du hamiltonien libre ($\mathbf{L} = 0$). C'est bien sûr une valeur moyenne quantique qu'il faudra injecter dans nos équations de Maxwell macroscopiques. Cet état est sans moment angulaire moyen et donc sans dipôle magnétique permanent (ce qui ne serait pas le cas d'un niveau P, par exemple). Il ne reste donc que la moyenne du terme en $\mathbf{A}$. En utilisant une des expressions possibles du potentiel vecteur d'un champ uniforme ($\frac{1}{2}\mathbf{B} \times \mathbf{R}$ et en notant que le champ magnétique, uniforme, n'est pas un opérateur, nous trouverons:

$$\langle \mathbf{m} \rangle = -\frac{q^2}{4m}\langle 1S|r^2\mathbf{B} - (\mathbf{R}.\mathbf{B})\mathbf{R}|1S\rangle \ . \tag{2.31}$$

$\mathbf{B}$ étant aligné avec Oz, cette expression se transforme instantanément en remarquant que $\langle 1S|ZX|1S\rangle = \langle 1S|ZY|1S\rangle = 0$:

$$\langle \mathbf{m} \rangle = -\frac{q^2 B}{4m}\langle 1S|X^2 + Y^2|1S\rangle \ . \tag{2.32}$$

On retrouve donc exactement, de manière assez surprenante, l'expression résultant du modèle naïf, où il suffit de remplacer r^2 par la valeur moyenne de $X^2 + Y^2$. Encore une fois, les modèles les plus simples se trouvent confirmés par une approche complètement quantique.

2.3.2 Paramagnétisme

Considérons maintenant le cas d'atomes ou de molécules portant un dipôle magnétique permanent (si des considérations d'invariance par renversement du sens du temps empêchent les atomes de posséder un moment dipolaire électrique permanent, ils peuvent porter un moment magnétique). Les molécules ayant un tel moment dipolaire, appelées paramagnétiques, ont en général un électron solitaire, non apparié sur son orbitale. C'est par exemple le cas de l'oxygène moléculaire, qui, sous sa forme liquide, est fortement attiré vers les zones de champs magnétiques forts. Comme pour les dipôles électriques, il doit exister une polarisabilité magnétique, positive, résultant de la compétition entre l'alignement des dipôles dans le champ magnétique et l'agitation thermique.

Si nous considérons le dipôle magnétique comme une quantité classique dont la projection sur un axe n'est pas quantifiée, nous pouvons traiter le problème par la statistique classique. La démarche est

absolument identique à celle utilisée pour la polarisabilité électrique d'orientation et nous ne donnons que le résultat final. En posant:

$$\xi = \frac{m_0 B_m}{kT} \ , \tag{2.33}$$

où m_0 est le module du dipôle moléculaire, on obtient:

$$\overline{m} = m_0 \left[\coth \xi - \frac{1}{\xi} \right] \ , \tag{2.34}$$

dont le développement pour des champs magnétiques faibles donne une polarisabilité d'orientation:

$$\alpha_o^m = \frac{\mu_0 m_0^2}{3kT} \ . \tag{2.35}$$

Comme il se devait, nous trouvons une polarisabilité positive. L'approximation de champs faibles est bien vérifiée pour tous les champs expérimentalement possibles et des températures voisines de l'ambiante. En effet, l'amplitude typique des dipôles magnétiques électroniques est le magnéton de Bohr, 14 GHz par Tesla en unités de fréquence (on exprime un moment magnétique en unités de fréquence en écrivant $mB = h\nu$). Une énergie thermique correspond, en fréquence, à une trentaine de THz. Les énergies magnétiques sont donc bien négligeables par rapport à l'agitation thermique. Notons que le magnétisme nucléaire, dû à l'orientation des moments magnétiques des noyaux, est encore plus faible, le moment magnétique d'un noyau étant typiquement 2000 fois plus faible que celui d'un électron. Bien sûr, ces ordres de grandeur ne tiennent que si on néglige les interactions entre moments magnétiques voisins par rapport au champ extérieur. En fait, dans les matériaux ferromagnétiques, une interaction d'origine purement quantique entre moments voisins peut, à une température suffisamment basse, aligner tous les spins du milieu et créer une aimantation intense. Là encore, nous ne nous préoccuperons pas de ce type de substance.

Comparons à ce point les ordres de grandeur des polarisabilités paramagnétiques et diamagnétiques, α_o et α_d. On a:

$$\frac{\alpha_o}{\alpha_d} = \frac{\mu_0 m_0^2 / 3kT}{q^2 a_0^2 \mu_0 / 4m} \ . \tag{2.36}$$

En remarquant que l'ordre de grandeur du moment magnétique permanent est

$$m_0 \simeq qc\alpha a_0 \ , \tag{2.37}$$

puisqu'il correspond à un électron sur une orbite de taille a_0 parcourue à la vitesse αc, où α est encore une fois la constante de structure fine. On a donc finalement:

$$\frac{\alpha_o}{\alpha_d} \simeq \frac{c^2 \alpha^2 m}{kT} \simeq \frac{R}{kT} \ . \tag{2.38}$$

Comme dans le cas électrique, le rapport des polarisabilités d'orientation et induite est de l'ordre de l'énergie de liaison atomique divisée par l'énergie thermique. On comprend donc bien que les susceptibilités paramagnétiques soient typiquement mille fois plus grandes que les susceptibilités diamagnétiques.

Cette approche classique, comme pour le dipôle électrique, est a priori insuffisante. En effet, le moment magnétique est, en mécanique quantique, à l'instar de toutes les observables vectorielles, proportionnel au moment cinétique. Sa projection sur l'axe Oz est donc quantifiée. Pour simplifier le calcul, nous supposerons que le moment angulaire de notre atome est de $\frac{1}{2}$. L'énergie d'interaction magnétique ne peut alors prendre que deux valeurs, $\omega_m I_z$, où $I_z = \pm\hbar/2$ est la projection sur z du moment cinétique atomique et $\omega_m = gB_m$ est la fréquence de Larmor, g étant le coefficient de proportionnalité entre le moment magnétique et le moment cinétique, appelé rapport gyromagnétique.

Comme seulement deux niveaux d'énergie sont accessibles, on peut écrire explicitement les probabilités d'occupation de ces niveaux:

$$n_{\pm} = \frac{\exp(\pm\hbar\omega_m/2kT)}{2\cosh\hbar\omega_m/2kT} \ . \tag{2.39}$$

L'aimantation moyenne est donnée simplement par $\overline{m} = (n_+ - n_-)g\hbar/2$. On trouve donc:

$$\overline{m} = \frac{g\hbar}{2}\tanh\frac{\hbar\omega_m}{2kT} \ . \tag{2.40}$$

En développant cette expression pour des champs magnétiques faibles, on trouve la polarisabilité quantique d'orientation:

$$\alpha^m_{o,q} = \frac{\mu_0\hbar^2 g^2}{4kT} \ . \tag{2.41}$$

A la définition près de la valeur du moment magnétique ($m_0 = g\hbar/2$), cette expression ne diffère du résultat classique que par un préfacteur numérique proche de l'unité. L'essentiel de la physique est donc contenu dans l'expression classique. Nous laissons au lecteur le soin d'établir l'expression de la polarisabilité pour un spin quelconque. Comme l'indique le principe de correspondance, on retrouvera l'expression classique dans la limite des très grands spins.

Notons encore une fois que nous n'avons établi que les expression statiques des polarisabilités. Elles sont cependant valables tant que la période des champs est longue par rapport aux constantes de temps d'atteinte de l'équilibre thermodynamique.

2.4 Lien entre polarisabilité et susceptibilité

Nous savons maintenant exprimer la polarisation des molécules en fonction du champ "perçu" par chaque molécule, $\mathbf{E}_m$ ou $\mathbf{B}_m$. Si le milieu était très peu dense, l'influence des molécules voisines serait négligeable et le champ vu par la molécule serait sensiblement identique au champ macroscopique. La polarisation par unité de volume serait alors simplement le produit du moment dipolaire de chaque molécule par la densité numérique de molécules (nombre de molécules par unité de volume). On aurait, pour la polarisation électrique, $\mathbf{P} = N\mathbf{p}_m = N\epsilon_0\alpha^e\mathbf{E}_m$. La susceptibilité serait alors simplement:

$$\chi^e = N\alpha^e \ , \tag{2.42}$$

manifestement sans dimension. Toutes les quantités écrites ci dessus se réfèrent bien sûr à des composantes monochromatiques. Pour alléger les écritures, nous négligerons souvent d'écrire explicitement les indices ω.

Si le calcul qui précède est correct pour des matériaux très dilués (les gaz par exemple), il ne s'applique pas dans la matière dense. En effet, dans ce cas, la molécule "voit" un champ qui est la somme d'une composante moyenne, proche du champ macroscopique, et du champ des molécules les plus proches. Ce champ est un champ microscopique, qu'il nous faudra représenter correctement pour obtenir la susceptibilité. Nous allons en fait, dans les deux prochains paragraphes, nous employer à montrer qu'il existe une relation linéaire entre le champ "vu" par la molécule et le champ macroscopique. Munis de cette relation, nous pourrons alors évaluer les susceptibilités en fonction des polarisabilités. Les cas électriques et magnétiques étant légèrement différents, nous les traiterons séparément.

2.4.1 Cas électrique

Le problème est donc de déterminer le champ "vrai", microscopique, "vu" par la particule m. Il s'agit d'un champ local, calculé au point précis où se trouve cette molécule, $\mathbf{r}_m$. Ce champ doit inclure le champ des sources macroscopiques et aussi le champ créé par toutes les particules du milieu

à l'exception toutefois de la particule m elle–même, qui ne peut contribuer à sa propre polarisation. Nous voyons poindre ici une différence importante avec le champ macroscopique, qui inclut les contributions de toutes les molécules *y compris* m et qui doit être évalué à un point quelconque du milieu.

Dans le champ $\mathbf{E}_m$, on peut distinguer deux types de contributions. D'abord celles des particules "lointaines", situées à une distance au moins de s_0, et ensuite la contribution des molécules "proches". Pour distinguer clairement entre ces deux contributions, nous isolerons les particules contenues dans une sphère de rayon voisin de s_0 centrée sur la molécule m. Il nous faut choisir une sphère pour respecter l'isotropie du système. Nous écrirons alors:

$$\mathbf{E}_m = \mathbf{E}_n + \mathbf{E}_f \ , \tag{2.43}$$

où l'indice n désigne la contribution des molécules proches et l'indice f celle des particules lointaines et des sources macroscopiques.

La contribution $\mathbf{E}_f$ des sources lointaines ne pose aucun problème. Elle est essentiellement constante sur l'ensemble du volume de la sphère. Pour calculer la contribution $\mathbf{E}_n$, nous pouvons d'abord négliger les phénomènes de propagation et raisonner comme en électrostatique. Nous sommes en effet, comme s_0 est beaucoup plus petit que la longueur d'onde (nous ne nous préoccupons bien sûr que d'une composante de fréquence), dans le domaine des champs proches, où le champ du dipôle est essentiellement le champ électrostatique. Bien sûr, la propagation devra être incluse soigneusement dans la contribution des molécules lointaines. $\mathbf{E}_n$ est donc la somme de tous les champs de tous les dipôles moléculaires autre que m. L'échelle s_0 étant très grande par rapport aux distances interatomiques, nous avons un très grand nombre de particules dans notre sphère. Nous ne ferons pas une trop grande erreur en calculant le champ $\mathbf{E}_n$ comme si toutes les molécules dans la sphère avaient toutes le même dipôle, égal au dipôle moyen à cet endroit, égal donc au dipôle que prendra la molécule m dans le champ total. L'opération rigoureuse serait de prendre les dipôles individuels et de moyenner les champs produits. Nous préférons ici prendre le dipôle moyen et évaluer le champ produit.

Nous noterons donc $\boldsymbol{\eta}(\boldsymbol{\rho})$ le champ produit au point $\mathbf{r}_m$ par un dipôle moyen situé au point $\mathbf{r}_m + \boldsymbol{\rho}$ (il est évident que ce champ ne dépend que de la distance relative de notre dipôle source et de la molécule de référence). Nous pouvons, en nous fondant sur le très grand nombre de particules dans la sphère, calculer le champ $\mathbf{E}_n$ par un argument statistique: c'est essentiellement l'intégrale sur la sphère du champ produit par une particule en $\boldsymbol{\rho}$ pondéré par la densité de probabilité $\mathcal{P}_c(\boldsymbol{\rho})$ de trouver une particule en ce point:

$$\mathbf{E}_n = \int \mathcal{P}_c(\boldsymbol{\rho})\boldsymbol{\eta}(\boldsymbol{\rho})\, d^3\boldsymbol{\rho} \ . \tag{2.44}$$

On prendra garde, dans cette expression, à ce que $\mathcal{P}_c$ est la densité de probabilité *conditionnelle* de trouver une particule à la position $\boldsymbol{\rho}$ par rapport au centre de la sphère en *sachant qu'il y a une molécule au centre de la sphère (m).* Cette probabilité est manifestement différente de celle de trouver une particule en un point quelconque de l'espace. Il est en particulier évident que $\mathcal{P}_c$ sera nulle si la distance ρ est plus petite que la "taille" des molécules. Finalement, le champ "vu" par la molécule m s'écrira:

$$\mathbf{E}_m = \mathbf{E}_f + \int \mathcal{P}_c(\boldsymbol{\rho})\boldsymbol{\eta}(\boldsymbol{\rho})\, d^3\boldsymbol{\rho} \ . \tag{2.45}$$

Nous pouvons tenir le même genre de raisonnement pour réévaluer le champ macroscopique. On distinguera aussi, pour le champ créé dans la sphère, la contribution des molécules lointaines et des molécules proches. La contribution lointaine, $\mathbf{E}_f$, est manifestement la même qu'auparavant. La différence essentielle est que nous devons estimer ce champ à un point arbitraire dans la sphère. La contribution des molécules proches peut dans ce cas s'écrire:

$$\int \mathcal{P}(\boldsymbol{\rho})\boldsymbol{\eta}(\boldsymbol{\rho})\, d^3\boldsymbol{\rho} \ , \tag{2.46}$$

en faisant intervenir cette fois la densité de probabilité $\mathcal{P}$ de trouver une molécule en un point arbitraire de la sphère (essentiellement la densité numérique pour un matériau homogène). En d'autres termes, pour estimer le champ vu par la molécule m, il nous faut exclure le champ de cette molécule elle-même, ce qui revient à ne considérer que la densité de probabilité conditionnelle, qui exclut explicitement le fait qu'une autre molécule puisse se trouver au centre de la sphère. Pour le champ ordinaire, on prend toutes les molécules, y compris m, et on n'exclut pas la probabilité qu'une molécule se trouve au point d'observation. On aura donc:

$$\mathbf{E} = \mathbf{E}_f + \int \mathcal{P}(\boldsymbol{\rho})\boldsymbol{\eta}(\boldsymbol{\rho})\, d^3\boldsymbol{\rho} \ . \tag{2.47}$$

En rapprochant les équations (2.45) et (2.47), nous voyons qu'on peut écrire le champ "vu" par la mOlécule en fonction du champ macroscopique et d'une différence d'intégrales qui reste à évaluer:

$$\mathbf{E}_m = \mathbf{E} + \int (\mathcal{P}_c(\boldsymbol{\rho}) - \mathcal{P}(\boldsymbol{\rho}))\boldsymbol{\eta}(\boldsymbol{\rho})\, d^3\boldsymbol{\rho} \ . \tag{2.48}$$

Evaluons d'abord l'intégrale sur $\mathcal{P}$. Pour un milieu isotrope, homogène à l'échelle macroscopique, $\mathcal{P}$ est à peu près uniforme et égale à la densité numérique N du milieu. L'intégrale de cette quantité est donc essentiellement le champ électrique créé par une sphère uniformément polarisée, avec une polarisation précisément égale à la polarisation macroscopique $\mathbf{P}$ (puisque chaque molécule de la sphère a le dipôle moyen). Le calcul de ce champ est un exercice d'électrostatique classique. On peut par exemple modéliser la situation par deux sphères uniformément chargées en volume, décalées spatialement d'une quantité très petite par rapport à leur rayon. On trouve alors que le champ intérieur est uniforme, avec une valeur

$$-\frac{\mathbf{P}}{3\epsilon_0} \ . \tag{2.49}$$

Examinons maintenant l'intégrale sur $\mathcal{P}_c$. Considérons d'abord le cas d'un milieu dense et désordonné comme un liquide. $\mathcal{P}_c$ doit alors avoir une symétrie sphérique par rapport à l'origine. Elle est certainement nulle à l'origine (les molécules ne peuvent s'interpénétrer). Elle doit ensuite croître et atteindre un maximum pour des distances de l'ordre de la distance intermoléculaire moyenne. On est en gros sûr de trouver une molécule à une distance moyenne d'une molécule donnée. Le comportement ensuite peut être complexe, mais, dans tous les cas, la probabilité conditionnelle tend rapidement vers la densité numérique. Au bout de quelques distances moyennes, les corrélations de position s'effacent dans les liquides et la probabilité conditionnelle tend vers la probabilité simple. L'intégrale sur $\mathcal{P}_c$ est donc essentiellement le champ créé en son centre par une distribution de polarisation à symétrie sphérique, pratiquement uniforme sauf un "trou" au voisinage du centre. Une application élémentaire du principe de superposition implique que ce champ est la somme des champs de deux sphères concentriques de polarisation opposées. Il est donc tout simplement nul.

Examinons maintenant le cas d'une structure cristalline. Considérons une structure cubique. Pour calculer le champ des plus proches voisins, nous avons à sommer 6 champs de dipôles, tous identiques, situés en $\pm a$ sur chacun des axes, a étant la maille cristalline. Supposons, pour simplifier, que tous les dipôles soient alignés avec Oz. Les quatre dipôles sur Ox et Oy contribuent à un champ antiparallèle à Oz, de module moitié et de direction opposée au champ produit par les dipôles selon Oz. Le champ résultant est donc strictement nul. En fait, pour toutes orientations des dipôles et toute structure cristalline, le champ résultant de l'intégrale sur $\mathcal{P}_c$ sera nul. On a donc, finalement:

$$\mathbf{E}_m = \mathbf{E} + \frac{\mathbf{P}}{3\epsilon_0} \ . \tag{2.50}$$

Le dipôle moyen sera donc:

$$\mathbf{p}_m = \epsilon_0 \alpha^e \mathbf{E}_m \ . \tag{2.51}$$

La densité macroscopique de polarisation s'écrit:

$$\mathbf{P} = N\mathbf{p}_m = \epsilon_0 \chi^e \mathbf{E} \ . \tag{2.52}$$

En rapprochant ces équations, on trouve enfin l'expression de la susceptibilité en fonction des polarisabilités:

$$\chi^e = \frac{N\alpha^e}{1 - N\alpha^e/3} \ , \tag{2.53}$$

expression valable quelle que soit la nature physique de la polarisabilité. Ce calcul a été établi pour la première fois en électrostatique par Clausius et Mosotti, et repris en régime quelconque, indépendamment, par Lorentz et Lorenz.

Pour des densités très faibles, on peut négliger le terme correctif au dénominateur et on trouve bien une susceptibilité $N\alpha$. En revanche, pour des milieux denses et des polarisabilités élevées, le terme correctif augmente la susceptibilité par rapport à l'extrapolation de celle d'un milieu dilué. Le champ créé par les molécules polarisées par le champ tend à renforcer cette polarisation. On peut estimer l'ordre de grandeur de cet effet. Les polarisabilités moyennes sont de l'ordre de a_0^3 (au voisinage des basses fréquences), les densités de quelques dixièmes de molécule par a_0^3 dans les milieux les plus denses. Le terme correctif est donc au plus de quelques dizaines de pour cent.

A priori toutefois, rien n'empêche qu'un milieu atteigne une densité telle que $N\alpha = 3$. La susceptibilité devrait alors diverger, les molécules s'orientant toutes seules sous l'influence de leur propre champ. Les matériaux ferroélectriques présentent ce genre de comportement, mais uniquement à fréquence nulle. On ne connaît aucun matériau qui puisse acquérir spontanément une polarisation alternative (hélas). On pourrait imaginer remplir ces conditions, par exemple, en partant d'un gaz peu dense et en l'examinant au voisinage d'une fréquence de résonance, par exemple de la vapeur de sodium sur la raie jaune. La polarisabilité d'un atome isolé est très grande, et on atteindrait facilement le seuil fatidique. En fait, ce type de raisonnement pêche en utilisant la polarisabilité d'un atome isolé. Quand la densité est grande, les atomes voisins se perturbent, par des interactions de van der Waals, par exemple, et leur polarisabilité diminue (une façon de le voir est que la largeur des raies atomiques augmente et que donc leur facteur de qualité diminue). En fait, la polarisabilité s'effondre très vite avec la densité et les facteurs correctifs ne dépassent jamais quelques dizaines de pour cent.

2.4.2 Cas magnétique

Il nous reste à traiter le cas magnétique. Le calcul est essentiellement le même, fondé sur la même comptabilité statistique des champs proches et lointains. On trouvera sans peine:

$$\mathbf{B}_m = \mathbf{B} + \int (\mathcal{P}_c(\boldsymbol{\rho}) - \mathcal{P}(\boldsymbol{\rho}))\boldsymbol{\beta}(\boldsymbol{\rho})\, d^3\boldsymbol{\rho} \ , \tag{2.54}$$

où les densités de probabilité ont les mêmes significations et où $\boldsymbol{\beta}$ est le champ magnétique créé à l'origine par le dipôle moyen situé en $\boldsymbol{\rho}$. On montrerait, comme précédemment que l'intégrale sur $\mathcal{P}_c$ est identiquement nulle (les géométries des cartes de champ du dipôle magnétique et du dipôle électrique sont identiques). Il ne nous reste donc qu'à évaluer le champ magnétique créé en son centre par une sphère uniformément magnétisée. Le calcul étant moins standard que pour le cas électrique, nous allons le traiter explicitement. Les équations de Maxwell macroscopiques dans la sphère s'écrivent:

$$\boldsymbol{\nabla} \cdot \mathbf{B} = 0 \tag{2.55}$$

$$\boldsymbol{\nabla} \times \mathbf{H} = 0 \ , \tag{2.56}$$

puisqu'il n'y a pas de courants libres. Il existe donc, dans ce cas, un potentiel magnétique scalaire Φ d'où dérive l'induction magnétique:

$$\mathbf{H} = -\boldsymbol{\nabla}\Phi \ . \tag{2.57}$$

Notons que ce potentiel scalaire magnétique n'existe que quand les courants macroscopiques sont nuls. Il ne permet de traiter que les matériaux aimantés, mais il est fort utile dans ce cas. Nous retrouvons

d'ailleurs ici la théorie des masses magnétiques, qui fut très utilisée au XIX siècle pour traiter des aimants. La nullité de la divergence de $\mathbf{B}$ implique que:

$$\boldsymbol{\nabla} \cdot \mathbf{H} + \boldsymbol{\mathcal{M}} = 0 \ , \tag{2.58}$$

où $\boldsymbol{\mathcal{M}}$ est l'aimantation. Le potentiel scalaire magnétique obéit donc à une équation de Poisson qui s'écrit:

$$\Delta\Phi = \boldsymbol{\nabla} \cdot \boldsymbol{\mathcal{M}} \tag{2.59}$$

Dans le cas de la sphère polarisée électriquement, la densité de charges liées est $\rho = -\boldsymbol{\nabla} \cdot \mathbf{P}$ et le potentiel scalaire obéit à l'équation de Poisson:

$$\Delta V = \frac{1}{\epsilon_0} \boldsymbol{\nabla} \cdot \mathbf{P} \ . \tag{2.60}$$

La similitude (à des facteurs dimensionnels près) des deux équations prouve que l'induction magnétique $\mathbf{H}$ a essentiellement la même forme que le champ électrique de la sphère polarisée:

$$\mathbf{H} = -\boldsymbol{\mathcal{M}}/3 \ , \tag{2.61}$$

d'où on déduit immédiatement:

$$\mathbf{B} = \mu_0 \frac{2}{3} \boldsymbol{\mathcal{M}} \ . \tag{2.62}$$

En injectant ce résultat dans le calcul du champ vu par la molécule, nous avons:

$$\mathbf{B}_m = \mathbf{B} - \mu_0 \frac{2}{3} \boldsymbol{\mathcal{M}} \ . \tag{2.63}$$

Le dipôle moyen induit est:

$$\mathbf{m}_m = \frac{1}{\mu_0} \alpha^m \mathbf{B}_m \ , \tag{2.64}$$

et l'aimantation moyenne est simplement:

$$\boldsymbol{\mathcal{M}} = N \mathbf{m}_m \ . \tag{2.65}$$

En rapprochant l'ensemble de ces résultats, on détermine la susceptibilité magnétique:

$$\chi^m = \frac{N\alpha^m}{1 + 2N\alpha^m/3} \ . \tag{2.66}$$

Pour la limite des faibles densités, on retrouve bien une susceptibilité qui est le produit de la polarisabilité par la densité numérique. Pour des densités plus fortes, le facteur correctif joue un rôle. Pour les substances paramagnétiques, la polarisabilité est positive. Le facteur correctif tend donc à réduire la susceptibilité. En fait, le champ des plus proches voisins est un champ démagnétisant, qui tend à s'opposer au champ macroscopique. La susceptibilité ne pose donc aucun problème de divergence. Pour des substances diamagnétiques, la polarisabilité est négative, et la susceptibilité se trouve renforcée. Toutefois, ces polarisabilités diamagnétiques sont en général très faibles et aucun effet de divergence n'est à craindre. Armés d'expressions réalistes pour les susceptibilités, nous pouvons maintenant les utiliser dans les équations de Maxwell macroscopiques.

2.5 Perméabilité et permittivité relatives

2.5.1 Définitions et équations de Maxwell

Nous avons donc:

$$\mathbf{D}_\omega = \epsilon_0(1+\chi^e)\mathbf{E}_\omega \ . \tag{2.67}$$

Nous ne manipulons forcément que des composantes monochromatiques, puisque les susceptibilités dépendent de la fréquence. Il est naturel de poser:

$$\epsilon_r(\mathbf{r},\omega) = 1+\chi^e \ . \tag{2.68}$$

Nous nommerons cette quantité "permittivité diélectrique relative". On a donc

$$\mathbf{D}_\omega = \epsilon_0\epsilon_r\mathbf{E}_\omega \tag{2.69}$$

Notons que ϵ_r s'exprime facilement en fonction de la polarisabilité:

$$\epsilon_r = \frac{1+2N\alpha^e/3}{1-N\alpha^e/3} \ . \tag{2.70}$$

Comme la polarisabilité électrique, induite ou d'orientation, est toujours positive à basse fréquence, la permittivité diélectrique relative est toujours supérieure à 1. En revanche, au dessus de la fréquence de résonance des système atomiques, la polarisabilité peut devenir négative et ϵ_r plus petit que 1. Notons que la relation précédente s'inverse facilement en:

$$N\alpha^e = 3\frac{\epsilon_r-1}{\epsilon_r+2} \ . \tag{2.71}$$

Cette relation, en électrostatique, est connue sous le nom de relation de Clausius–Mosotti.

Nous pouvons utiliser le même genre d'arguments pour les phénomènes magnétiques. On peut écrire:

$$\mathbf{B}_\omega = \mu_0\mu_r(\mathbf{r},\omega)\mathbf{H}_\omega \ , \tag{2.72}$$

avec

$$\mu_r = \frac{1}{1-\chi^m} \ , \tag{2.73}$$

que nous appellerons "perméabilité magnétique relative". Là encore, il ne s'agit que de relations entre composantes de Fourier. μ_r est donc supérieur à un pour les matériaux paramagnétiques, de polarisabilité ou de susceptibilité positive, inférieur à un pour les matériaux diamagnétiques. On peut aussi écrire μ_r en fonction de la polarisabilité sous la forme:

$$\mu_r = \frac{1+2N\alpha^m/3}{1-N\alpha^m/3} \ , \tag{2.74}$$

équation strictement identique à celle que nous avions écrite pour ϵ_r. La relation inverse est aussi identique à la relation de Clausius–Mosotti électrique:

$$N\alpha^m = 3\frac{\mu_r-1}{\mu_r+2} \ . \tag{2.75}$$

Nous n'avons ici que des relations entre composantes de Fourier. Dans le cas général, les équations de Maxwell ne prennent une forme simple qu'après une transformée de Fourier temporelle. **D** et **H** s'exprimant en fonction des champs électriques et magnétiques, on obtient un ensemble de quatre équations portant sur deux champs de vecteurs seulement qu'il est, au moins en principe, possible de résoudre. Dans le cas particulier important où les susceptibilités sont pratiquement indépendantes de

la fréquence dans tout le domaine de fréquence couvert par la physique du problème, il est possible de regrouper les composantes de Fourier et d'écrire les équations de Maxwell comme:

$$\begin{aligned}
\boldsymbol{\nabla} \times \mathbf{E} &= -\frac{\partial \mathbf{B}}{\partial t} && (2.76)\\
\boldsymbol{\nabla} \cdot \mathbf{B} &= 0 && (2.77)\\
\boldsymbol{\nabla} \cdot (\epsilon_r \mathbf{E}) &= \rho/\epsilon_0 && (2.78)\\
\boldsymbol{\nabla} \times (\mathbf{B}/\mu_r) &= \mu_0 \left[\mathbf{j} + \epsilon_0 \frac{\partial \epsilon_r \mathbf{E}}{\partial t}\right] . && (2.79)
\end{aligned}$$

Nous avons bien pris garde de ne pas sortir les permittivités et perméabilités relatives des dérivations. Elles peuvent en effet a priori dépendre de l'espace, si le matériau n'est pas homogène, et du temps. Ce n'est que dans un matériau homogène et invariable qu'on retrouve une forme identique à celle des équations de Maxwell dans le vide, avec les simples substitutions $\epsilon_0 \to \epsilon_0\epsilon_r$ et $\mu_0 \to \mu_0\mu_r$.

2.5.2 Considérations énergétiques

Nous pouvons maintenant reprendre l'équation bilan énergétique. Nous nous restreindrons à un milieu dont les susceptibilités sont indépendantes du temps. Le terme de variation d'énergie électromagnétique peut s'écrire:

$$\mathbf{E} \cdot \frac{\partial \mathbf{D}}{\partial t} + \mathbf{H} \cdot \frac{\partial \mathbf{B}}{\partial t} = \frac{\partial u}{\partial t} \tag{2.80}$$

à condition de poser:

$$u = \frac{\epsilon_0\epsilon_r}{2} E^2 + \frac{B^2}{2\mu_0\mu_r} . \tag{2.81}$$

De son côté, le vecteur de Poynting s'écrit:

$$\mathbf{\Pi} = \frac{\mathbf{E} \times \mathbf{B}}{\mu_0\mu_r} . \tag{2.82}$$

on retrouve donc strictement les bilans énergétiques de l'espace libre, avec les substitutions $\epsilon_0 \to \epsilon_0\epsilon_r$ et $\mu_0 \to \mu_0\mu_r$.

Nous avions vu dans le chapitre précédent que la puissance cédée par les champs à la matière pour faire varier sa polarisation s'exprimait comme:

$$\mathbf{E} \cdot \frac{\partial \mathbf{P}}{\partial t} . \tag{2.83}$$

Bien sûr, il y a aussi une puissance fournie pour modifier les polarisations magnétiques. Toutefois, la plupart des matériaux avec des propriétés magnétiques marquées sont opaques et ne présentent guère d'intérêt du point de vue de la propagation. Nous ne considérerons donc plus qu'un matériau dit diélectrique, équivalent au vide du point de vue magnétique ($\mu_r = 1$) et ne présentant que des propriétés électriques. Si nous ne considérons qu'un champ monochromatique, avec des amplitudes $\mathbf{E}_0$ et $\mathbf{P}_0$ pour le champ électrique et la polarisation, la valeur moyenne temporelle de cette puissance s'écrira:

$$\frac{1}{2}\mathrm{Re}\,(-i\omega\mathbf{P}_0) \cdot \mathbf{E}_0^*) . \tag{2.84}$$

En utilisant la susceptibilité $\chi = \chi' + i\chi''$ (on omettra la mention ^{e}. Nous ne discutons que du cas électrique et il n'y a pas de risque de confusion avec une susceptibilité magnétique), on met cette puissance sous la forme:

$$\frac{1}{2}|E_0|^2\epsilon_0\omega\chi'' . \tag{2.85}$$

La puissance moyenne cédée par le champ à la matière est donc proportionnelle à la partie imaginaire de la susceptibilité. Cette puissance reflète bien sûr l'absorption d'une onde électromagnétique par la matière. Si χ'' est positif, le transfert d'énergie se produit du champ vers la matière: cette puissance reflète bien sûr l'absorption d'une onde électromagnétique par la matière. Si χ'', en revanche, est négatif, la matière fournit de l'énergie à l'onde. On peut alors s'attendre à ce qu'une onde électromagnétique soit amplifiée par un tel milieu. On se convaincra aisément que la susceptibilité résultant du modèle de Thomson de l'électron élastiquement lié a une partie imaginaire positive. Ce modèle correspond toujours à une absorption, l'énergie fournie par l'onde servant à compenser l'énergie dissipée par le frottement phénoménologique introduit dans l'équation du mouvement de l'électron.

Le modèle quantique, pour sa part, permet une situation un peu plus riche. Si nous prenons le modèle quantique tel que nous l'avons formulé dans la partie précédente, la polarisabilité est strictement réelle, et aucune absorption ne peut se produire. Nous avons en effet un modèle qui néglige complètement la dissipation atomique. Nous pouvons facilement rendre le modèle plus réaliste en rajoutant phénoménologiquement un terme d'amortissement dans l'expression de la polarisabilité ou de la susceptibilité. Pour un milieu suffisamment dilué pour que les facteurs correctifs des équations de Clausius–Mosotti soient négligeables, la susceptibilité s'écrira donc:

$$\chi = \frac{Nq^2}{2m\epsilon_0} \sum_j \frac{f_{jg}}{\omega_{jg}^2 - \omega^2 - i\gamma_{jg}\omega} \ , \tag{2.86}$$

où la somme est à étendre à tous les niveaux j autres que le fondamental, f_{jg} étant la force d'oscillateur de la transition de g vers j et finalement γ_{jg} un facteur d'amortissement phénoménologique pour cette transition (on trouve ce facteur de façon exacte en quantifiant le champ et en tenant compte correctement de l'émission spontanée). On peut donc extraire facilement la partie imaginaire de la susceptibilité. Au voisinage immédiat de la résonance, le domaine dans lequel il peut se passer des choses intéressantes, on a:

$$\chi'' = \frac{Nq^2}{2m\epsilon_0} \sum_j \frac{\gamma_{jg}}{\omega} \frac{f_{jg}}{4(\omega_{jg} - \omega)^2 + \gamma_{jg}^2} \ . \tag{2.87}$$

On trouve donc qu'au voisinage d'une résonance, la partie imaginaire de χ et donc l'échange d'énergie avec le champ présente un comportement Lorentzien. De plus, toutes les quantités étant positives à part la force d'oscillateur, χ'' a le signe de f_{jg}. Celui-ci est le signe de la fréquence de Bohr ω_{jg} de la transition résonante. Rappelons que, dans le calcul quantique de la susceptibilité, nous avons supposé que pratiquement toute la population atomique est dans le niveau g. Si donc le niveau j a une énergie supérieure au niveau g, χ'' est positif et le milieu absorbe l'énergie du champ comme dans le modèle classique de Thomson. En revanche, si le niveau j se trouve énergétiquement au dessous du niveau le plus peuplé, g, la partie imaginaire de la susceptibilité devient négative. Dans ce cas, le milieu atomique fournit de l'énergie au champ et l'amplifie.

Nous trouvons ici en fait une version semi–classique de l'émission stimulée: sur une transition inversée, où la population du niveau supérieur est plus grande que la population du niveau du bas, les photons émis par les atomes sous l'influence d'une onde incidente, s'ajoutent de façon cohérente à l'onde incidente et l'amplifient. Un tel milieu amplificateur couplé à un résonateur peut conduire à une oscillation permanente. En un mot, nous venons d'inventer le laser! On peut comprendre assez bien le fonctionnement d'un laser en remplissant un interféromètre de Fabry Perot avec un milieu amplificateur. Dans un Fabry Perot ordinaire, la finesse finie est due aux pertes subies par le faisceau lumineux sur un aller et retour dans la cavité. Si le gain du milieu amplificateur est juste égal aux pertes, la finesse diverge et il peut exister un champ permanent dans la cavité. Si le gain est supérieur aux pertes, l'intensité croit exponentiellement. En fait, elle ne croît ainsi que sur une gamme limitée. La saturation de la transition atomique, que nous avions brièvement évoquée dans la partie précédente, fait que le gain est en fait une fonction décroissante de l'intensité. Il s'établit alors un régime stationnaire avec une intensité importante dans la cavité.

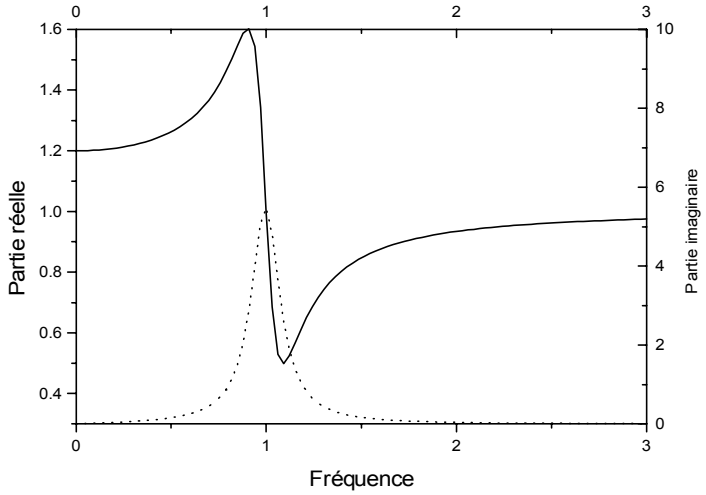

Figure 2.2: Partie réelle et imaginaire de la susceptibilité au voisinage de la résonance. On a représenté en traits pleins la courbe de dispersion $\mathrm{Re}\,(1+\chi)$ et en pointillés la courbe d'absorption $\mathrm{Im}\,\chi$. La fréquence est en unités de la fréquence de résonance. Le facteur de qualité de la résonance est de 30, très inférieur à ce qu'on a pour une résonance atomique.

Nous montrerons dans un paragraphe suivant que la partie réelle de la susceptibilité est, pour sa part, liée à l'indice de réfraction ordinaire. La susceptibilité contient donc tout à la fois les phénomènes de réfraction ou de dispersion, si l'indice n'est pas indépendant de la fréquence, et les phénomènes d'absorption. On vérifiera sans peine que la partie réelle de la susceptibilité présente, voisinage d'une résonance atomique, une forme voisine de la dérivée d'une Lorentzienne. Elle est nulle strictement à résonance, maximale ou minimale environ à une largeur de la résonance stricte. Une telle courbe est appelée courbe de dispersion. Elle décrit de façon assez réaliste la variation de l'indice de réfraction ordinaire au voisinage de la résonance. La variation de l'indice avec la longueur d'onde étant liée aux phénomènes de dispersion, le nom a été conservé pour cette forme de courbes. Des courbes d'absorption et de dispersion typiques sont représentées sur la figure 2.2

2.6 Relations de Kramers–Krönig

On pourrait penser à priori que dispersion et absorption sont deux phénomènes indépendants. On pourrait alors imaginer par exemple des milieux qui, dans une très large gamme de fréquence, auraient une très forte absorption en dépit d'un indice voisin de un. Ce seraient de parfaits pièges à lumière. On pourrait aussi imaginer des verres qui feraient la fortune des lunetiers en étant de très grand indice et néanmoins totalement transparents, là aussi dans une large gamme de fréquences. En fait, nous allons montrer que la simple causalité (la polarisation ne peut répondre au champ avant que celui–ci n'ait été appliqué) impose des relations très fortes entre les parties réelles et imaginaires de l'indice. En fait, quand la première est connue sur tout le spectre, la seconde peut aisément être déterminée par une simple transformation mathématique.

Le point de départ est le lien entre polarisation et champ par la fonction de Green, transformée de Fourier de la susceptibilité:

$$\mathbf{P}(t) = \frac{\epsilon_0}{\sqrt{2\pi}} \int G(\tau)\mathbf{E}(t-\tau) \ . \tag{2.88}$$

toutes ces quantités s'entendant à un point $\mathbf{r}$ donné. En fait, nous ferons le calcul pour la susceptibilité électrique, mais il s'appliquerait aussi bien à toutes nos susceptibilités, ou à toute forme de réponse

linéaire. Les relations que nous allons établir sont donc très générales, dépassant largement le cadre de l'électromagnétisme. Nous considérons ici la susceptibilité comme scalaire. Des aménagements de détail du calcul permettraient de traiter le cas tensoriel. Nous savons que la fonction de Green est la transformée de Fourier de la susceptibilité:

$$G(\tau) = \frac{1}{\sqrt{2\pi}} \int \chi(\omega) e^{-i\omega\tau}\, d\omega \ . \tag{2.89}$$

Cette intégrale peut bien sûr être transformée en une intégrale sur le plan complexe. Le chemin d'intégration est constitué de l'axe réel, parcouru dans le sens positif et fermé par un demi-cercle "à l'infini" (on pourra se reporter à la figure du chapitre sur les potentiels retardés, troisième partie). Si $\tau < 0$, le demi–cercle doit se trouver dans le demi–plan supérieur. C'est à cette condition que l'exponentielle $e^{-i\omega\tau}$ annule l'intégrale sur ce demi–cercle. En revanche, pour $\tau > 0$ on doit boucler le contour d'intégration dans le demi–plan inférieur. La causalité impose bien sûr que $G(\tau)$ s'annule pour $\tau < 0$. L'intégrale sur le contour dans le demi–plan supérieur est donc identiquement nulle. Cela impose, l'exponentielle étant régulière sur tout le plan complexe, que la susceptibilité χ *ne présente de pôles que dans le demi–plan inférieur*. On pourra vérifier aisément que les susceptibilités déduites du modèle de Thomson ou du modèle quantique présentent cette propriété. Notons que des pôles situés strictement sur l'axe réel correspondraient à des résonances sans amortissement (voir le calcul quantique de la quatrième partie), ce qui n'est guère physique.

Considérons maintenant la fonction:

$$\frac{\chi(\omega)}{2i\pi(\omega - z_0)} \ , \tag{2.90}$$

où z_0 a une partie imaginaire positive. Cette fonction admet donc un pôle unique en z_0 dans le demi–plan supérieur, dont le résidu est trivialement $\chi(z_0)/2i\pi$. L'intégrale sur le contour composé de l'axe réel et du demi–cercle supérieur est alors égale à ce résidu (multiplié par $2i\pi$). Pour tout système physique, la susceptibilité doit tendre rapidement vers 0 quand le module de la fréquence tend vers l'infini. L'intégrale sur le "demi–cercle à l'infini" est donc négligeable, et l'intégrale sur le contour se ramène à celle le long de l'axe réel. On a donc finalement:

$$\chi(z_0) = \int_{-\infty}^{\infty} \frac{\chi(\omega)}{2i\pi(\omega - z_0)} \ . \tag{2.91}$$

Posons maintenant:

$$z_0 = \omega_0 + i\epsilon \ , \tag{2.92}$$

avec ω_0 réel et prenons la limite de l'expression précédente pour $\epsilon \to 0$. Le point z_0 tend alors vers le point ω_0 de l'axe réel. Le premier membre, puisque χ est régulière dans tout le demi–plan supérieur, tend vers $\chi(\omega_0)$. On a donc:

$$\chi(\omega_0) = \frac{1}{2i\pi} \lim_{\epsilon\to 0} \int \frac{\chi(\omega)}{\omega - \omega_0 - i\epsilon} \ . \tag{2.93}$$

On montre facilement, en théorie des distributions, que [3]:

$$\lim_{\epsilon\to 0} \int \frac{1}{\omega - \omega_0 - i\epsilon} = \mathcal{PP}(1/(\omega - \omega_0)) + i\pi\delta(\omega - \omega_0) \ . \tag{2.94}$$

[3] On peut comprendre qualitativement cette limite en observant que la fonction à intégrer est, pour sa partie réelle, une courbe de dispersion (dérivée de Lorentzienne) et, pour sa partie imaginaire, une simple Lorentzienne. Ces deux courbes sont de largeur ϵ. Quand $\epsilon \to 0$, la seconde tend trivialement vers une fonction de Dirac. La première estime la différence entre deux valeurs infiniment proches de la fonction sur laquelle porte la distribution, ce qui est bien, fondamentalement, l'action de la distribution partie principale.

La distribution "partie principale en ω_0", $\mathcal{PP}$ est, rappelons le, définie par:

$$\left\langle \mathcal{PP}\frac{1}{\omega-\omega_0}|f\right\rangle = \lim_{\eta\to 0}\left[\int_{-\infty}^{\omega_0-\eta} + \int_{\omega_0+\eta}^{\infty} \frac{f(\omega)}{\omega-\omega_0}\right] \tag{2.95}$$

On a donc finalement, en constatant que l'action de δ sur χ est triviale,

$$\chi(\omega_0) = \frac{1}{i\pi}\mathcal{PP}\int \frac{\chi(\omega)}{\omega-\omega_0}\ . \tag{2.96}$$

En isolant, dans cette expression, les parties réelles et imaginaires, on obtient enfin les relations de Kramers–Krönig proprement dites:

$$\chi'(\omega_0) = \frac{1}{\pi}\mathcal{PP}\int \frac{\chi''(\omega)}{\omega-\omega_0} \tag{2.97}$$

$$\chi''(\omega_0) = -\frac{1}{\pi}\mathcal{PP}\int \frac{\chi'(\omega)}{\omega-\omega_0}\ . \tag{2.98}$$

Les parties réelle et imaginaire de la susceptibilité sont donc reliées par une transformation mathématique très simple. On peut comprendre ainsi pourquoi, si la partie imaginaire au voisinage d'une résonance, a un comportement Lorentzien, la partie réelle a la forme d'une courbe de dispersion. La partie principale calcule en effet une "sorte de dérivée" de la fonction. Elle est nulle si la fonction est paire par rapport au point ω_0, non nulle quand la fonction est impaire. La partie réelle est donc nulle au maximum de la partie imaginaire, au sommet de la résonance. Elle est en revanche importante quand la partie imaginaire varie rapidement, dans les ailes de la résonance. On voit donc bien que ces parties réelle et imaginaire sont fortement liées et que l'une ne saurait être grande sans que l'autre ne soit également importante, dans un domaine de fréquence voisin. Ces relations ont de nombreuses conséquences en théorie de la réponse linéaire. Elles peuvent, en pratique, être utilisées pour déterminer la réfraction si seule l'absorption est expérimentalement disponible (c'est par exemple le cas pour les milieux fortement absorbants), ou l'inverse.

Chapitre 3

Propagation dans les milieux linéaires

Nous allons maintenant appliquer les équations de Maxwell macroscopiques et les susceptibilités au problème de la propagation de champs monochromatiques dans des milieux matériels. Nous pourrons préciser le contenu physique de la dispersion et de l'absorption. Nous nous préoccuperons en fait essentiellement de deux types de milieux: les diélectriques transparents et les milieux conducteurs, regroupant métaux et plasmas. Dans un deuxième temps, nous établirons les lois de continuité des champs au voisinage d'une interface entre deux matériaux différents. Nous appliquerons ces lois au passage d'une onde plane entre deux matériaux diélectriques. Nous en déduirons les lois de Descartes de l'optique géométrique, mais aussi les coefficients de Fresnel donnant les amplitudes relatives des ondes transmises et réfléchies.

3.1 Equations de propagation

Nous considérerons donc la propagation d'un champ harmonique, de fréquence ω dans un milieu matériel. Ce milieu est caractérisé par les permittivité diélectrique ϵ_r, perméabilité magnétique μ_r relatives et par la conductivité σ. Nous supposerons le milieu homogène et ces quantités indépendantes de la position. Nous supposerons que le courant macroscopique ne résulte que de la conductivité du matériau et du champ électrique (en d'autres termes, nous supposons que les générateurs sont en dehors de notre milieu). Les équations de Maxwell en régime harmonique s'écrivent alors:

$$\begin{aligned}
\boldsymbol{\nabla} \times \mathbf{E} &= i\omega \mathbf{B} && (3.1)\\
\boldsymbol{\nabla} \cdot \mathbf{B} &= 0 && (3.2)\\
\boldsymbol{\nabla} \cdot \mathbf{E} &= \frac{\rho}{\epsilon_0 \epsilon_r} && (3.3)\\
\boldsymbol{\nabla} \times \mathbf{B} &= \mu_0 \mu_r \left[\sigma - i\epsilon_0 \epsilon_r \omega\right] \mathbf{E} \ . && (3.4)
\end{aligned}$$

A ces équations, nous pouvons adjoindre l'équation de conservation de la charge (qui se déduit des équations de Maxwell), qui s'écrit, avec $\mathbf{j} = \sigma \mathbf{E}$:

$$\sigma \boldsymbol{\nabla} \cdot \mathbf{E} = i\omega\rho \ . \tag{3.5}$$

En y injectant l'équation de Gauss, on trouve:

$$i\omega\rho = \sigma\rho/\epsilon_0\epsilon_r \ , \tag{3.6}$$

qui n'admet comme solution que $\rho = 0$ à moins que

$$\sigma = i\omega\epsilon_0\epsilon_r \ . \tag{3.7}$$

Ce cas étant a priori plutôt rare, nous l'exclurons pour le moment et nous pourrons donc supposer que ρ est toujours identiquement nulle dans les milieux matériels.

Munis de cette simplification considérable, nous pouvons facilement écrire une équation de propagation pour le champ électrique seul ou le champ magnétique seul. En fait, comme dans l'espace libre, les équations de propagation de tous les champs et de tous les potentiels sont identiques. Cette équation peut se mettre sous la forme d'une équation de Helmholtz:

$$\Delta \mathbf{E} + \frac{\Omega^2}{c^2}\mathbf{E} = 0 \ , \tag{3.8}$$

où la fréquence (complexe) Ω est définie par:

$$\Omega^2 = (\omega^2 \mu_0 \mu_r \epsilon_0 \epsilon_r + \mu_0 \mu_r i\omega\sigma)c^2 \ . \tag{3.9}$$

Dans le vide, on retrouve bien sûr $\Omega = \omega$. Toute la physique de la propagation est contenue dans cette fréquence. La solution en termes d'ondes planes avec une direction de propagation selon Oz:

$$\mathbf{E}_0 e^{i(kz-\omega t)} \tag{3.10}$$

devra vérifier la relation de dispersion

$$k^2 = \Omega^2/c^2 \ . \tag{3.11}$$

3.2 Milieux diélectriques

Nous considérerons d'abord le plus simple des matériaux, un matériau diélectrique ($\epsilon_r \neq 1$), dépourvu de toutes propriétés magnétiques ($\mu_r = 1$) et strictement isolant ($\sigma = 0$). La relation de dispersion s'écrit alors:

$$k = \frac{\omega}{c}\sqrt{\epsilon_r} \ . \tag{3.12}$$

Nous poserons:

$$n = \sqrt{\epsilon_r} \ , \tag{3.13}$$

et écrirons donc la relation de dispersion sous la forme:

$$k = n\frac{\omega}{c} = nk_0 \ , \tag{3.14}$$

où $k_0 = \omega/c$ serait le module du vecteur d'onde pour une propagation dans le vide à cette fréquence. Nous appellerons n l'indice de réfraction de notre matériau diélectrique. Il donne en effet le rapport entre la longueur d'onde dans le vide et la longueur d'onde dans le milieu.

Comme n est en général une quantité complexe, ainsi que ϵ_r ou χ^e, l'indice de réfraction est aussi une quantité complexe:

$$n = n' + in'' \ . \tag{3.15}$$

On peut montrer simplement que:

$$n' = \frac{1}{\sqrt{2}}\sqrt{\epsilon_r' + \sqrt{{\epsilon_r'}^2 + \epsilon_r''^2}} \tag{3.16}$$

et que

$$n'' = \frac{\epsilon_r''}{\sqrt{2}}\frac{1}{\sqrt{\epsilon_r' + \sqrt{{\epsilon_r'}^2 + \epsilon_r''^2}}} \ . \tag{3.17}$$

Notons que, pour respecter le sens choisi pour la propagation de l'onde, le signe de n' doit être positif. Il n'y a donc pas d'ambiguïté sur le signe de n''. Comme $\epsilon_r'' = \chi''$, la partie imaginaire de l'indice est

liée directement au coefficient d'absorption du milieu. On peut en effet remarquer que l'expression de l'onde plane de vecteur d'onde complexe peut se réécrire:

$$\mathbf{E}_0 e^{-n'' k_0 z} e^{i(k_0 n' z - \omega t)} \ . \tag{3.18}$$

Il s'agit d'une onde dont l'amplitude décroît exponentiellement lors de la propagation dans le milieu. La partie réelle, pour sa part, correspond à une augmentation du module du vecteur d'onde, donc à une réduction de la longueur d'onde (si la partie réelle de l'indice est plus grande que un, ce qui est fréquemment le cas). Il s'agit donc de l'indice de réfraction tel que le conçoit l'optique géométrique. La vitesse de phase de cette onde évanescente (la vitesse à laquelle se propagent les plans d'égale phase ou plans d'onde) est manifestement donnée par:

$$v_\phi = \frac{c}{n'} \ . \tag{3.19}$$

Elle est donc plus petite que la vitesse de la lumière dans un facteur égal à la partie réelle de l'indice de réfraction. La vitesse de groupe, pour sa part, dépend en général de la dépendance de n avec ω. On établira sans peine que:

$$v_g = \frac{c/n}{1 + \frac{\omega}{n}\frac{dn}{d\omega}} \ . \tag{3.20}$$

Au voisinage d'une résonance atomique, l'indice de réfraction varie rapidement avec ω et peut être une fonction décroissante de ω. On peut avoir, dans cette région de "dispersion anormale" une vitesse de groupe très petite ou très grande devant c, voire même une vitesse de groupe négative. C'est bien sûr la notion même de vitesse de groupe qui perd son sens dans ce cas. Elle ne s'applique a priori qu'à un paquet d'onde suffisamment large en fréquence pour être bien localisé dans l'espace, mais très étroit devant l'échelle de variation de la vitesse de propagation pour se propager encore sans déformation notable. Au voisinage d'une résonance très étroite, ces conditions ne sont pas remplies et le paquet se déforme considérablement, laissant une très grande marge à l'interprétation dans la définition de sa vitesse de propagation. Un certain nombre d'annonces spectaculaires récentes de "propagation supraluminique" reposent sur cette ambiguïté. En fait, quelle que soit la forme du paquet et de la dispersion, on montre qu'aucun signal ne peut se propager plus vite que c.

Nous avons ici les expressions exactes des parties réelles et imaginaires de l'indice de réfraction. Elles se simplifient beaucoup si le milieu et peu dense et/ou la susceptibilité, réelle comme imaginaire, est très inférieure à un. On peut en effet dans ce cas développer les expressions précédentes et obtenir:

$$n' = 1 + \chi'/2 \simeq 1 + N\alpha'/2 \tag{3.21}$$

$$n'' = \chi''/2 \simeq N\alpha''/2 \ . \tag{3.22}$$

Ces expressions de l'indice en fonction de la polarisabilité sont très utiles en pratique.

L'équation de propagation ne précise pas les amplitudes $\mathbf{E}_0$ et $\mathbf{B}_0$ des champs électriques et magnétiques. Leur divergence étant nulle, ils sont tous deux perpendiculaires au vecteur d'onde k. La première équation de Maxwell s'écrit, pour une onde plane de vecteur d'onde (complexe) $\mathbf{k}$:

$$i\mathbf{k} \times \mathbf{E}_0 = i\omega \mathbf{B}_0 \ . \tag{3.23}$$

On en déduit que la structure de l'onde plane est conservée, $\mathbf{E}_0$, $\mathbf{B}_0$ et $\mathbf{k}$ formant encore un trièdre direct. En revanche, pour un indice complexe, les champs ne sont pas nécessairement en phase, et le rapport de leurs modules est $B_0/E_0 = |n|/c$.

Dans le cas d'un milieu absorbant, l'énergie de l'onde passe à la matière sur une distance de l'ordre de $1/n''k_0$. Des notions comme le vecteur de Poynting ont donc un intérêt physique limité. Pour établir des bilans énergétiques, nous allons donc considérer le cas d'un milieu diélectrique complètement transparent ($n'' = 0$). Notons que les relations de Kramers Kronig n'interdisent nullement à l'absorption

d'être nulle pour une fréquence avec une partie réelle d'indice non nulle à la même fréquence. Les champs électriques et magnétiques sont alors en phase, et on a: $B_0 = nE_0/c$. Le vecteur de Poynting, sans surprises, pointe donc dans la direction du vecteur d'onde, et sa valeur moyenne temporelle s'écrit:

$$\overline{\Pi} = \frac{\epsilon_0 nc}{2} E_0^2 \ . \tag{3.24}$$

La densité d'énergie électromagnétique s'écrit, en valeur instantanée:

$$u = \epsilon_0 n^2 \frac{E^2}{2} + \frac{1}{2\mu_0} B^2 \ . \tag{3.25}$$

En utilisant le rapport entre les amplitudes de $\mathbf{E}$ et $\mathbf{B}$, on montre que, comme pour l'onde plane dans le vide, les densités d'énergies électriques et magnétiques sont égales à chaque instant. La valeur moyenne temporelle de la densité d'énergie s'écrit finalement:

$$\overline{u} = \epsilon_0 n^2 \frac{E_0^2}{2} \ . \tag{3.26}$$

En faisant maintenant le rapport de la valeur moyenne du vecteur de Poynting avec la valeur moyenne de la densité d'énergie, on doit obtenir, comme pour l'onde plane dans le vide, la vitesse de propagation de l'énergie, qui doit coïncider avec la vitesse de groupe[1]:

$$v_g = \frac{\overline{\Pi}}{\overline{u}} = \frac{c}{n} \ . \tag{3.27}$$

La vitesse de groupe, ou de propagation de l'énergie, est donc égale à la vitesse de phase, et plus petite (si $n > 1$) que la vitesse de la lumière. Il s'agit bien sûr d'un résultat très classique d'optique géométrique.

3.3 Milieux conducteurs

Nous allons maintenant considérer la propagation dans un milieu conducteur, pourvu de propriétés diélectriques, mais dépourvu de propriétés magnétiques. Nous aurons donc $\sigma \neq 0$, $\epsilon_r \neq 1$, mais $\mu_r = 1$. Nous supposerons de plus que la permittivité diélectrique relative est pratiquement indépendante de la fréquence. On pourra ainsi traiter de la propagation dans les métaux. Les électrons libres sont alors responsables de la conductivité, en général élevée, alors que les ions métalliques sont responsables des propriétés diélectriques. Les électrons libres, dans un tel matériau, sont mis en mouvement globalement par tout champ appliqué et contribuent donc au courant macroscopique, libre, alors que les autres électrons et ions, fixes, ne contribuent qu'aux charges liées. Nous pourrons aussi traiter des plasmas, où électrons et ions libres contribuent tous deux aux courants libres. Cependant, la masse des ions est toujours très grande par rapport à celle des électrons. Pour des champs de haute fréquence, les ions ne suivent pratiquement pas le champ appliqué et on peut considérer que seuls les électrons participent à la conduction et aux courants libres, les ions restant spectateurs et ne contribuant qu'aux propriétés diélectriques par leur polarisabilité. En d'autres termes, au prix de cette approximation, on voit que les propagations dans un plasma ou dans un métal sont essentiellement identiques.

3.3.1 Modèle de conductivité

Avant de décrire en détails la propagation, nous allons donner un modèle simple mais réaliste de conductivité, qui nous permettra de préciser sa dépendance en fréquence. Ce modèle est très universel

[1] Nous faisons ici en fait, de façon implicite, l'hypothèse que le spectre en fréquence du paquet d'ondes que l'on propage est très étroit par rapport au domaine de variation de l'indice avec la fréquence. Si l'indice ne dépend pas de la fréquence, $v_g = d\omega/dk$ est précisément égal à c/n.

et s'applique aussi bien aux électrolytes qu'aux plasmas ou aux métaux. Nous ne discuterons toutefois les ordres de grandeur que pour un métal bon conducteur, comme le cuivre. Nous ne considérerons aussi qu'un seul type de porteurs de charges libres que nous assimilerons à des électrons. C'est le cas pour les métaux et une bonne approximation pour les plasmas.

Sous l'action du champ électrique $\mathbf{E}$, les porteurs libres se mettent en mouvement d'ensemble. En général, ce mouvement d'ensemble se superpose aux mouvements aléatoires d'origine thermique (pour les plasmas) ou quantiques (principe d'exclusion de Fermi pour les métaux). Bien que ces vitesses aléatoires soient beaucoup plus grandes que les vitesses d'ensemble[2], nous allons écrire uniquement une équation sur la vitesse d'ensemble $\mathbf{v}$:

$$m\dot{\mathbf{v}} + m\gamma\mathbf{v} = q\mathbf{E} \ , \tag{3.28}$$

où m et q sont la masse et la charge des porteurs, et où γ est un facteur phénoménologique de friction, décrivant l'inévitable amortissement de la vitesse d'ensemble sous l'influence des collisions avec le réseau. La solution évidente de cette équation en régime harmonique est:

$$\mathbf{v} = \frac{q\mathbf{E}}{m(\gamma - i\omega)} \ . \tag{3.29}$$

Le courant macroscopique étant donné par:

$$\mathbf{j} = Nq\mathbf{v} \ , \tag{3.30}$$

où N est la densité numérique de porteurs au point considéré, on peut finalement écrire la conductivité comme:

$$\sigma = \frac{Nq^2}{m(\gamma - i\omega)} = \frac{\omega_p^2 \epsilon_0 \epsilon_r}{\gamma - i\omega} \ . \tag{3.31}$$

Nous avons introduit la "fréquence de plasma" du conducteur (dont la signification physique apparaîtra plus clairement dans les prochains chapitres):

$$\omega_p^2 = \frac{Nq^2}{m\epsilon_0\epsilon_r} \ . \tag{3.32}$$

Nous trouvons donc en général une conductivité complexe. Le mouvement des porteurs de charges a en effet un temps de réponse au champ appliqué, de l'ordre de γ^{-1}. Le courant ne suivant pas instantanément le champ, il est déphasé par rapport à celui-ci. Ce n'est que pour des fréquences très petites devant γ qu'on retrouve une conductivité réelle.

Pour fixer un peu les ordres de grandeur, considérons un bon métal (cuivre) à très basse fréquence. La conductivité est de l'ordre de $10^7\ \Omega^{-1}\mathrm{m}^{-1}$. La densité numérique d'électrons s'obtient facilement en calculant le nombre d'atomes par unité de volume (chacun participant par un électron à la conduction). On a environ $N = 10^{28}\ \mathrm{m}^{-3}$. En assimilant la permittivité diélectrique relative à 1, on trouve ainsi l'ordre de grandeur du taux de relaxation $\gamma \simeq 10^{13}\mathrm{s}^{-1}$. Cet ordre de grandeur est relativement réaliste. Pour une propagation d'électrons dans un réseau métallique, il correspond assez bien au temps moyen entre deux collisions avec les défauts du réseau (joints de grains, dislocations....). Avec ces mêmes ordres de grandeur, la fréquence plasma et de 10^{16} Hz. Elle est très élevé par rapport au taux de relaxation. Les longueurs d'onde correspondant au taux de relaxation sont dans le domaine de l'infrarouge lointain alors que celles correspondant à la fréquence de plasma sont dans l'ultraviolet proche.

[2]L'ordre de grandeur de la vitesse d'ensemble dans les métaux est le millimètre par seconde, alors que les vitesses de Fermi sont de l'ordre du centième de celle de la lumière.

Avec ce modèle simple de la conductivité, on peut écrire le vecteur d'onde complexe d'une onde plane se propageant dans le milieu en fonction de ω (relation de dispersion). On obtiendra:

$$k^2 = k_0^2 \epsilon_r \left[1 + i \frac{\omega_p^2}{\omega(\gamma - i\omega)} \right] . \tag{3.33}$$

En général, k présente une partie imaginaire et l'onde est absorbée dans le milieu. Les électrons, soumis à une force de friction, dissipent en effet de l'énergie. Pour préciser un peu le comportement de k, on peut séparer trois domaines de fréquence selon les valeurs relatives de ω, ω_p et γ.

3.3.2 Propagation très basse fréquence

Nous considèrerons ici une fréquence très basse par rapport au taux de relaxation:

$$\omega \ll \gamma \ll \omega_p . \tag{3.34}$$

La conductivité est alors pratiquement réelle:

$$\sigma \simeq \frac{Nq^2}{m\gamma} = \frac{\omega_p^2 \epsilon_0 \epsilon_r}{\gamma} , \tag{3.35}$$

(on retrouve le modèle de conductivité de Drude dans ce régime basse fréquence) et on peut écrire:

$$k^2 \simeq k_0^2 \epsilon_r \left[1 + i \frac{\omega_p^2}{\gamma\omega} \right] . \tag{3.36}$$

Dans l'expression entre crochets, le second terme est très grand devant 1, qui peut être négligé. L'extraction de la racine ne pose aucun problème et on trouve:

$$k = \pm(1+i)/\delta , \tag{3.37}$$

avec

$$\delta = \frac{1}{k_0} \sqrt{\frac{2\gamma\omega}{\epsilon_r \omega_p^2}} = c \sqrt{\frac{2\epsilon_0}{\sigma\omega}} , \tag{3.38}$$

où on a noté que $\omega_p^2 = \sigma\gamma/\epsilon_0\epsilon_r$. Les parties réelles et imaginaires du vecteur d'onde sont donc égales. L'onde s'amortit sur une distance δ, égale à la longueur d'onde dans le milieu. Une onde électromagnétique ne peut pénétrer dans un milieu conducteur. C'est l'effet de peau, bien connu des techniciens des hautes fréquences.

Précisons l'ordre de grandeur de la profondeur de peau, δ, pour un bon métal comme le cuivre. Elle est proportionnelle à la racine carrée de la longueur d'onde dans le vide à la même fréquence. Pour une très basse fréquence, $\omega = 2\pi \times 50$ Hz, on trouve $\delta \simeq 3$ cm. La profondeur de peau à une fréquence aussi basse est en général très grande par rapport à l'épaisseur des circuits ou des fils. On néglige cet effet le plus souvent en électrotechnique sauf sur les lignes à très haute puissance, de grand diamètre en général. Cet ordre de grandeur montre aussi qu'il est très difficile d'écranter efficacement un circuit électronique des très basses fréquences. Pour une fréquence beaucoup plus élevée, $2\pi \times 50$ Mrd/s, par exemple, la profondeur de peau n'est plus que de 30 μm. A haute fréquence, le champ ne pénètre pratiquement pas dans le conducteur. Quand on veut transmettre une puissance importante, il est donc préférable de donner au conducteur la forme d'un ruban très aplati, ayant une épaisseur de l'ordre de la profondeur de peau. Cet ordre de grandeur prouve aussi que l'on peut très efficacement écranter un circuit électronique du bruit haute fréquence en le plaçant dans un blindage métallique d'épaisseur modérée.

3.3.3 Propagation haute fréquence

Considérons maintenant le cas où la fréquence est supérieure à la fréquence de plasma et donc très grande par rapport au taux d'amortissement:

$$\omega > \omega_p \gg \gamma \ . \tag{3.39}$$

La relation de dispersion s'écrit alors, en négligeant γ:

$$k^2 = k_0^2 \epsilon_r \left[1 - \frac{\omega_p^2}{\omega^2} \right] \ . \tag{3.40}$$

La quantité dans le crochet étant réelle positive, la propagation s'effectue avec un vecteur d'onde réel, sans amortissement (nous avons en effet éliminé le taux de relaxation). L'indice de réfraction est:

$$n = \sqrt{\epsilon_r \left(1 - \frac{\omega_p^2}{\omega^2} \right)} \ . \tag{3.41}$$

Dans le cas où $\epsilon_r = 1$, cet indice est plus petit que l'unité. Cela reflète le fait que la vitesse de phase de l'onde:

$$v_\phi = \frac{\omega}{k} = \frac{c}{n} = \frac{c}{\sqrt{\epsilon_r}} \frac{1}{\sqrt{1 - \omega_p^2/\omega^2}} \ , \tag{3.42}$$

est plus grande que la vitesse de la lumière. Cette vitesse de phase ne correspond pas à la vitesse de propagation de l'énergie, donnée par la vitesse de groupe. Rien ne lui impose donc d'être plus petite que c. La vitesse de groupe est donnée par:

$$v_g = \frac{d\omega}{dk} = \frac{c}{\sqrt{\epsilon_r}} \sqrt{1 - \omega_p^2/\omega^2} \ . \tag{3.43}$$

Elle est donc plus petite que $c/\sqrt{\epsilon_r}$, comme il se doit. On remarquera d'ailleurs que:

$$v_g v_\phi = \frac{c^2}{\epsilon_r} \ . \tag{3.44}$$

Cette relation, assez générale, est également vérifiée pour la propagation dans les guides d'onde, la fréquence de plasma étant alors remplacée par la fréquence de coupure du guide. Notons que la vitesse de groupe tend vers zéro et que la vitesse de phase diverge quand la fréquence tend vers la fréquence de plasma. Au voisinage immédiat de cette fréquence, l'onde ne se propage pratiquement plus. La fréquence de plasma joue donc bien le rôle d'une fréquence de coupure dans ce problème. Le comportement en fréquence de la vitesse de groupe et de la vitesse de phase est illustré sur la figure 3.1.

3.3.4 Régime intermédiaire

Nous considérerons maintenant les fréquences plus faibles que la fréquence de plasma mais néanmoins très supérieures au taux d'amortissement:

$$\omega_p > \omega \gg \gamma \ . \tag{3.45}$$

La relation de dispersion peut alors encore s'écrire, en négligeant γ:

$$k^2 = k_0^2 \epsilon_r \left[1 - \frac{\omega_p^2}{\omega^2} \right] \ . \tag{3.46}$$

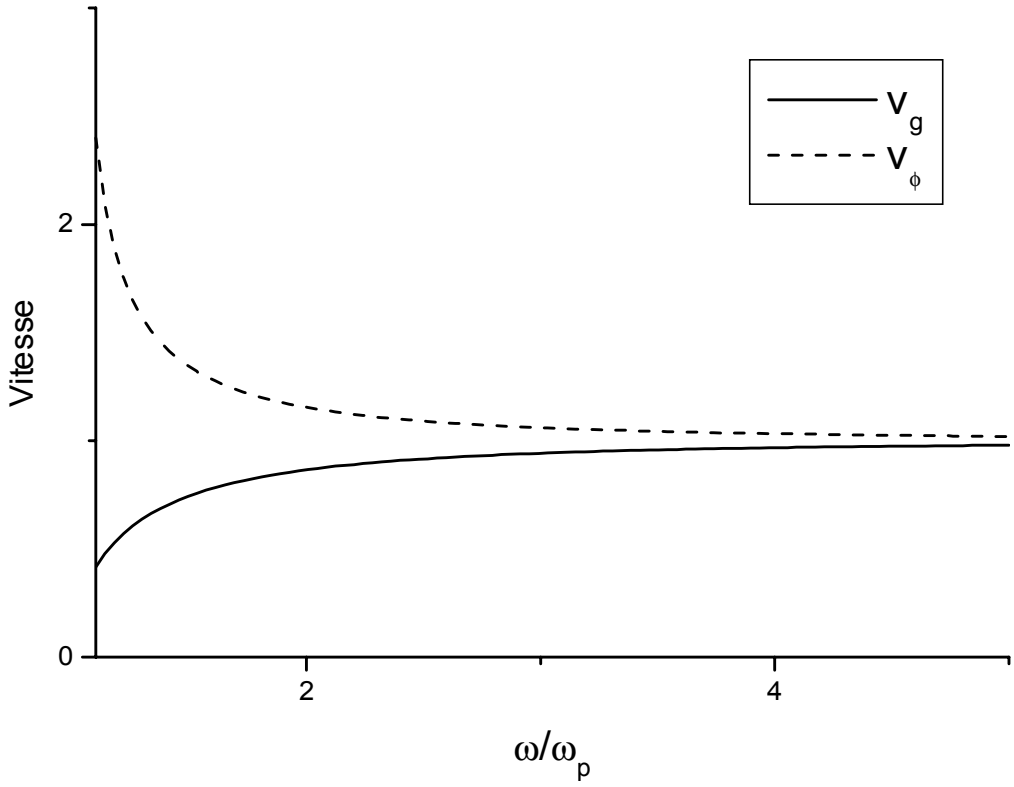

Figure 3.1: Vitesse de groupe et vitesse de phase en fonction de ω/ω_p, en unités de la vitesse de la lumière dans le milieu.

La quantité entre crochets est maintenant négative et k est un imaginaire pur. Nous avons donc une onde dont la phase est partout la même (vitesse de phase infinie) et dont l'amplitude décroît exponentiellement dans le milieu avec une profondeur de pénétration:

$$\delta = \frac{c}{\sqrt{\epsilon_r(\omega_p^2 - \omega^2)}} \ . \tag{3.47}$$

L'ordre de grandeur de cette profondeur de peau, sauf au voisinage immédiat de la fréquence plasma où elle diverge, est la longueur d'onde de plasma, c/ω_p. Il s'agit donc d'une quantité très inférieur au micron. Il y a donc, à la fréquence plasma, une discontinuité très nette entre deux régimes: un régime de très forte absorption au dessous de la fréquence plasma et un régime de transparence parfaite au dessus. Le conducteur doit donc présenter un comportement résonnant à la fréquence plasma. C'est ce comportement que nous allons explorer maintenant.

3.3.5 Ondes de Plasma

Si nous nous plaçons à la fréquence de plasma, nous pouvons négliger le taux d'amortissement γ. Nous pouvons alors voir que la conductivité vérifie:

$$\sigma = i\epsilon_0\epsilon_r\omega_p \ . \tag{3.48}$$

En établissant les équations de propagation, nous avions explicitement exclu ce cas. C'est en effet seulement quand cette condition est vérifiée qu'il peut exister une densité de charges dans le milieu. Si ρ est non nulle, il peut exister un champ électrique de divergence non nulle. Si nous cherchons ce champ sous la forme d'une onde plane, il s'agit d'une onde longitudinale, dont l'amplitude est colinéaire au vecteur d'onde. Nous l'appellerons onde de plasma. En prenant la direction de propagation selon Oz, nous pouvons donc chercher une solution en onde plane sous la forme:

$$E_z e^{i(kz-\omega t)}\mathbf{u}_z \ . \tag{3.49}$$

On peut alors écrire le rotationnel de $\mathbf{B}$:

$$\boldsymbol{\nabla} \times \mathbf{B} = \mu_0[\sigma - i\omega_p\epsilon_0\epsilon_r]\mathbf{E} \ . \tag{3.50}$$

Le crochet étant précisément nul à la fréquence de plasma, le rotationnel de $\mathbf{B}$ est nul. Le champ magnétique, dont divergence et rotationnel sont nuls, est donc identiquement nul.

Physiquement, l'onde que nous venons de décrire correspond donc à une modulation sinusoïdale de la densité de porteurs libres dans l'espace à un instant donné. Il en résulte un champ électrique longitudinal qui tend à ramener les charges à leur position d'équilibre. En l'absence d'amortissement, le plasma oscille autour de cette position d'équilibre, à une fréquence déterminée seulement par sa densité, ω_p. On pourra se convaincre aisément que le vecteur d'onde de l'onde de plasma est complètement arbitraire car il n'existe aucune relation de dispersion.

Dans un plasma assez dense, les champs électriques associés aux ondes de plasmas peuvent être très intenses. Il a été suggéré de les utiliser pour réaliser des accélérateurs de particules. Un battement entre deux faisceaux lasers de fréquences supérieures à la fréquence de plasma, se propageant donc librement, se couple par des effets non linéaires à la densité de charges et excite une onde de plasma. Il est possible, en choisissant correctement les vecteurs d'onde incidents d'accorder la vitesse de phase de l'onde de plasma à celle de la particule en accélération et de faire en sorte qu'elle soit toujours soumise à un champ électrique accélérateur. Au moins sur le papier, ces accélérateurs linéaires à plasma devraient avoir des performances tout à fait remarquables. Il ne reste qu'à vaincre les nombreuses difficultés techniques.

3.4 Relations de passage

Nous avons considéré jusque-là la propagation dans un milieu linéaire homogène. Nous allons maintenant, en route vers l'optique géométrique, nous poser le problème d'une discontinuité plane (au moins localement) entre deux milieux linéaires de caractéristiques différentes. Nous aurons à l'esprit surtout les interfaces entre diélectriques transparents mais la plupart de nos résultats s'appliqueront à des milieux arbitraires. Nous commencerons par établir les relations de continuité vérifiées par les champs électrique et magnétique de part et d'autre de l'interface. Nous en déduirons ensuite, en considérant le passage d'une onde plane d'un milieu à l'autre, les lois de Descartes de la réflexion et de la réfraction. Enfin, nous établirons les relations de Fresnel qui lient les amplitudes des ondes réfléchie, réfractée et incidente.

3.4.1 Conditions aux limites

Nous considérons donc une interface entre deux milieux 1 et 2, au voisinage d'un point où la normale à l'interface, orientée conventionnellement de 1 vers 2, est $\mathbf{n}$. Nous calculerons flux et circulation des champs sur deux types d'objets.

Nous considérerons d'abord une "boîte de Camembert, cylindrique, de base dS. Les deux faces sont parallèles à la surface, situées immédiatement au dessus pour l'une et au dessous pour l'autre. La paroi latérale a une surface infinitésimale par rapport à dS. Le flux de $\mathbf{B}$ à travers la boîte doit être identiquement nul. Le flux sur la surface latérale étant négligeable, on a donc:

$$(\mathbf{B}_2 - \mathbf{B}_1) \cdot \mathbf{n} = 0 \ , \tag{3.51}$$

où $\mathbf{B}_2$ et $\mathbf{B}_1$ désignent les champs magnétiques de part et d'autre de la surface, au voisinage immédiat du point considéré. La composante normale à la surface du champ magnétique est donc continue. En appliquant le même raisonnement au déplacement électrique, on a:

$$(\mathbf{D}_2 - \mathbf{D}_1) \cdot \mathbf{n} = \sigma \ , \tag{3.52}$$

où σ est une éventuelle densité surfacique de charges libres localisée sur la surface. La composante normale du déplacement électrique présente donc la même discontinuité qu'en électrostatique.

Les deux autres équations de Maxwell portent sur des rotationnels. Nous utiliserons donc la circulation des champs sur un contour rectangulaire. Les deux grands côtés sont parallèles à la surface, de longueur dl selon un vecteur unitaire $\mathbf{u}$. Un des côtés est situé dans le milieu 1, l'autre dans le milieu 2. Les deux petits côtés, perpendiculaires à l'interface, ferment le rectangle. Leur longueur est infinitésimale par rapport à dl. La circulation de $\mathbf{E}$ sur ce contour est égale à la dérivée temporelle du flux du champ magnétique à travers le contour. Le champ magnétique est sûrement borné au voisinage de l'interface, même s'il est discontinu. Le flux de ce champ dans le contour, dont la hauteur est infinitésimale par rapport à la longueur, est donc négligeable. Le circulation de $\mathbf{E}$ sur ce contour est donc nulle:

$$(\mathbf{E}_2 - \mathbf{E}_1) \cdot \mathbf{u} dl = 0 \ . \tag{3.53}$$

$\mathbf{u}$ étant un vecteur arbitraire perpendiculaire à $\mathbf{n}$, cette relation exprime que la composante tangentielle du champ électrique est continue. Elle peut aussi s'écrire:

$$(\mathbf{E}_2 - \mathbf{E}_1) \times \mathbf{n} = 0 \ . \tag{3.54}$$

Ecrivons enfin l'équation de Maxwell–Ampère. La circulation de $\mathbf{H}$ sur le contour fait intervenir la dérivée par rapport au temps du flux de $\mathbf{D}$. Comme $\mathbf{B}$ plus haut, $\mathbf{D}$ est une quantité bornée au voisinage de l'interface et ce flux est négligeable. En revanche, le courant de conduction des charges libres peut avoir un flux non nul à travers le contour s'il s'agit d'un courant surfacique. Ce flux pourra s'écrire $dl\, \mathbf{j}_s \cdot \mathbf{v}$, en posant $\mathbf{v} = \mathbf{n} \times \mathbf{u}$ (il s'agit évidemment du vecteur définissant la normale à la

surface rectangulaire s'appuyant sur notre contour). $\mathbf{j}_s$ est la densité de courant surfacique, intégrale d'un courant volumique très localisé au voisinage de l'interface, sur une épaisseur infinitésimale par rapport à toutes les longueurs considérées ici. Notons que la dimension de $\mathbf{j}_s$ est celle d'un courant divisé par une longueur (ou d'une densité de courant multipliée par une longueur). On obtient donc une intensité en intégrant $\mathbf{j}_s$ sur une ligne tracée sur la surface. On a donc finalement:

$$(\mathbf{H}_2 - \mathbf{H}_1) \cdot \mathbf{u} = \mathbf{j}_s \cdot \mathbf{v} = \mathbf{j}_s \cdot (\mathbf{n} \times \mathbf{u}) = \mathbf{u} \cdot (\mathbf{j}_s \times \mathbf{n}) \ . \tag{3.55}$$

Cette relation étant vraie pour toute orientation de $\mathbf{u}$ parallèlement à la surface, on a:

$$\mathbf{H}_2 - \mathbf{H}_1 = \mathbf{j}_s \times \mathbf{n} \ , \tag{3.56}$$

ou:

$$\mathbf{n} \times (\mathbf{H}_2 - \mathbf{H}_1) = \mathbf{j}_s \ . \tag{3.57}$$

On peut finalement regrouper les conditions au limites pour les quatre champs sous la forme:

$$\mathbf{n} \cdot (\mathbf{D}_2 - \mathbf{D}_1) = \sigma \tag{3.58}$$
$$\mathbf{n} \times (\mathbf{E}_2 - \mathbf{E}_1) = 0 \tag{3.59}$$
$$\mathbf{n} \cdot (\mathbf{B}_2 - \mathbf{B}_1) = 0 \tag{3.60}$$
$$\mathbf{n} \times (\mathbf{H}_2 - \mathbf{H}_1) = \mathbf{j}_s \ , \tag{3.61}$$

pour lesquelles on prendra garde que $\mathbf{n}$ est orienté de 1 vers 2. En ajoutant les relations définissant les susceptibilités, ces équations déterminent complètement les discontinuités des champs au passage de la surface. On aura évidemment noté que ces équations sont les mêmes qu'en électrostatique ou en magnétostatique. Comme nous n'avons, pour les établir, considéré qu'un domaine infinitésimal au voisinage de la surface, il est assez clair que les phénomènes de propagation, ou même d'induction, ne peuvent jouer aucun rôle.

3.4.2 Passage entre deux milieux diélectriques

Considérons maintenant, pour nous rapprocher encore de l'optique, une onde plane incidente sur un dioptre plan entre deux milieux diélectriques, non magnétiques, de permittivités diélectriques relatives (éventuellement complexes) ϵ_1 et ϵ_2. Nous noterons $\mathbf{k}_i$ le vecteur d'onde incident. Cette onde se propage dans le milieu 1. L'interface étant dépourvue de toute densité de charges ou de courants libres, les composantes tangentielles de $\mathbf{E}$ et de $\mathbf{H}$ et les composantes normales de $\mathbf{D}$ et de $\mathbf{B}$ sont continues. En fait, les relations de dispersion dans les deux milieux et ces conditions aux limites ne peuvent être satisfaites que s'il existe, dans le milieu 2, une onde transmise de vecteur d'onde différent $\mathbf{k}_t$ et une onde réfléchie dans le milieu 1, de vecteur d'onde $\mathbf{k}_r$.

Lois de Descartes

Les conditions de passage entre les deux milieux étant linéaires, elles ne pourront être satisfaites à tous les instants que si ces trois ondes ont exactement la même fréquence ω. Notons qu'il existe des interfaces, parfois très simples comme la surface de l'eau, qui possèdent des propriétés non linéaires et sont capables de réfléchir l'harmonique deux de la fréquence incidente. Ce genre de phénomène, fournissant de précieux renseignements sur la structure de l'interface (orientation des molécules par exemple) étant très marginal, nous le négligerons ici.

Toutes les relations de passage peuvent donc, en éliminant les facteurs oscillants communs, se mettre sous la forme:

$$a e^{i\mathbf{k}_i \cdot \mathbf{r}} + b e^{i\mathbf{k}_r \cdot \mathbf{r}} + c e^{i\mathbf{k}_t \cdot \mathbf{r}} = 0 \ , \tag{3.62}$$

où $\mathbf{r}$ est un point du dioptre et a, b, c des coefficients constants ne dépendant que de la géométrie des ondes et de leurs polarisations. Pour que ces relations soient vérifiées quel que soit $\mathbf{r}$, il faut évidemment que $(\mathbf{k}_i - \mathbf{k}_r)\cdot\mathbf{r}$ et $(\mathbf{k}_i - \mathbf{k}_t)\cdot\mathbf{r}$ soient des constantes en fonction de $\mathbf{r}$. En prenant l'origine sur la surface, ces constantes sont nulles et on en déduit que $\mathbf{k}_i - \mathbf{k}_r$ et $\mathbf{k}_i - \mathbf{k}_t$ sont perpendiculaires à la surface. On en déduit donc la première loi de Descartes: les trois vecteurs d'onde ainsi que la normale $\mathbf{n}$ au dioptre sont dans un même plan.

Dans ce plan, nous définirons l'angle d'incidence θ_i comme l'angle de $\mathbf{k}_i$ avec la normale au dioptre orientée dans la direction du rayon incident ($-\mathbf{n}$ avec nos notations). Nous définirons de même l'angle de réflexion, θ_r et l'angle de réfraction, θ_t, défini, cette fois, par rapport à $\mathbf{n}$. On peut alors écrire, dans le plan d'incidence, l'égalité des composantes selon la surface des trois vecteurs d'onde. Quand on note que les modules de $\mathbf{k}_i$ et $\mathbf{k}_r$ sont les mêmes puisque ces ondes se propagent dans le même milieu, on en déduit immédiatement l'égalité des angles d'incidence et de réflexion (deuxième loi de Descartes). Finalement, en écrivant que $k_i = n_1\omega/c$ et $k_t = n_2\omega/c$, où n_1 et n_2 sont les indices de réfraction des deux milieux, on établit:

$$n_1 \sin\theta_i = n_2 \sin\theta_t \ , \tag{3.63}$$

la troisième loi de Descartes. Finalement, les lois de la réflexion et de la réfraction de l'optique géométrique se déduisent simplement de l'existence et de la linéarité des conditions aux limites. Elles sont donc très générales et peuvent, par exemple, s'appliquer avec des modifications mineures à l'acoustique.

La troisième loi de Descartes contient bien sûr le phénomène de réfraction limite quand on passe d'un milieu d'indice fort à un milieu d'indice faible. Il n'y a plus alors d'angle réel qui satisfasse aux lois de Descartes pour le faisceau transmis et toute l'énergie incidente est réfléchie. On peut cependant accorder dans ce cas un peu plus de crédit aux lois de Descartes qu'on ne le fait habituellement. L'angle de réfraction devrait avoir un sinus plus grand que 1. C'est donc un angle complexe. Le vecteur d'onde dans le milieu 2 est donc un vecteur d'onde complexe (au moins dans la direction normale). On a donc, dans le milieu 2, une onde évanescente, avec une profondeur de peau de l'ordre de la longueur d'onde (divisée par 2π quand on fait le calcul complet). Nous laissons au lecteur le soin de déterminer plus précisément les caractéristiques de cette onde.

Coefficients de Fresnel

Nous n'avons utilisé, pour établir les lois de Descartes, que l'existence des conditions de passage et leur linéarité. En fait, ces relations sont connues et nous pouvons maintenant les utiliser pour écrire explicitement les relations entre les amplitudes des ondes incidente et réfléchie. La géométrie du problème est assez complexe si nous considérons une onde incidente de polarisation quelconque. Nous pouvons utiliser une fois de plus la linéarité et ne considérer en fait que deux cas: une polarisation dans le plan d'incidence (polarisation conventionnellement appelée π) et une polarisation perpendiculaire au plan d'incidence (appelée polarisation σ). Les relations de continuité des composantes imposent bien sûr que, si le champ électrique incident est π (σ) les champs réfléchis et transmis seront aussi de polarisation π (σ). Les polarisations des différentes ondes sont donc équivalentes. Les conventions d'orientation pour les amplitudes des champs électrique et magnétique sont résumées sur la figure 3.2.

Considérons d'abord le cas de la polarisation σ. Parmi toutes les relations possibles, nous écrirons la continuité des composantes tangentielles de $\mathbf{E}$ et $\mathbf{B}$. Ces deux relations s'écrivent, en termes des amplitudes complexes, avec les conventions d'orientation de la figure:

$$E_i + E_r = E_t \tag{3.64}$$

$$-B_i \cos\theta_i + B_r \cos\theta_r = -B_t \cos\theta_t \ . \tag{3.65}$$

En remarquant que $B = nE/c$, on peut écrire la seconde relation en fonction des champs électriques. En exprimant alors $E_r = E_t - E_i$ à partir de la première relation et en utilisant l'égalité des angles de

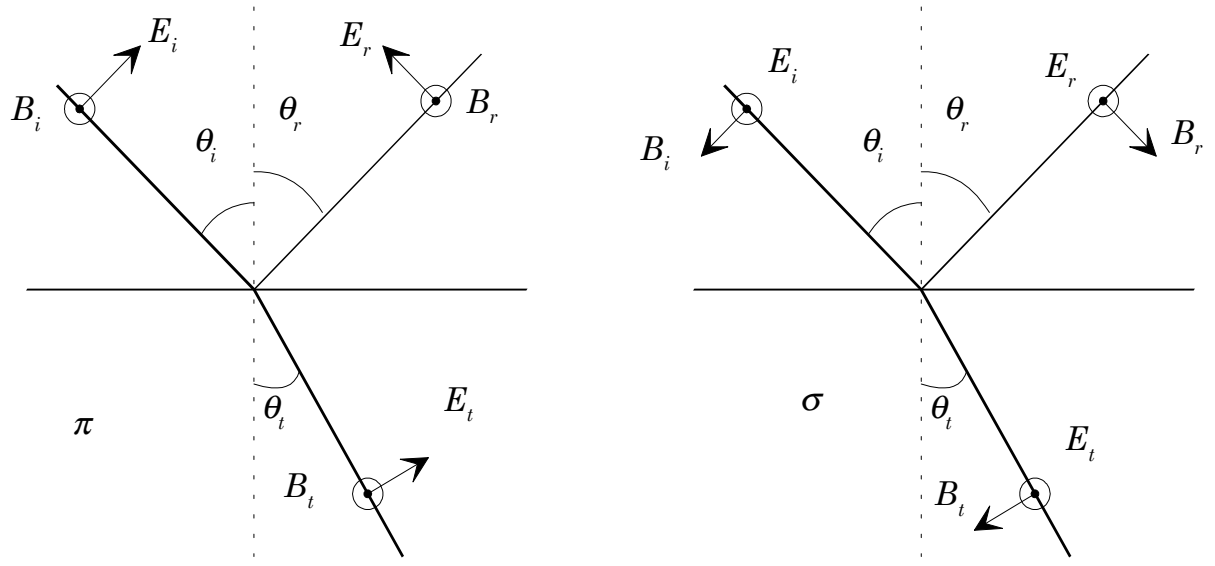

Figure 3.2: Conventions d'orientation des champs électrique et magnétique pour l'établissement des relations de Fresnel. A gauche, polarisation dans le plan d'incidence (π). A droite, polarisation perpendiculaire au plan d'incidence (σ)

réflexion et de réfraction, on tire de la seconde égalité:

$$E_t = \frac{2n_1 \cos\theta_i}{n_2 \cos\theta_t + n_1 \cos\theta_i} E_i \tag{3.66}$$

et enfin

$$E_r = \frac{n_1 \cos\theta_i - n_2 \cos\theta_t}{n_2 \cos\theta_t + n_1 \cos\theta_i} E_i \ . \tag{3.67}$$

Ce sont les relations de Fresnel qui donnent les amplitudes des ondes réfléchie et réfractée en fonction de celle de l'onde incidente.

En polarisation π, le calcul est essentiellement le même. On écrit les continuités des composantes tangentielles de **E** et **B**:

$$B_t = B_i + B_r \tag{3.68}$$

$$E_i \cos\theta_i - E_r \cos\theta_r = E_t \cos\theta_t \ . \tag{3.69}$$

On en déduit les relations de Fresnel:

$$E_t = \frac{2n_1 \cos\theta_i}{n_1 \cos\theta_t + n_2 \cos\theta_i} E_i \tag{3.70}$$

$$E_r = \frac{n_2 \cos\theta_i - n_1 \cos\theta_t}{n_1 \cos\theta_t + n_2 \cos\theta_i} E_i \tag{3.71}$$

Le cas de l'incidence normale mérite une considération particulière. En polarisation σ, en prenant tous les angles nuls, on trouve:

$$E_t = \frac{2n_1}{n_1 + n_2} E_i \tag{3.72}$$

$$E_r = \frac{n_1 - n_2}{n_1 + n_2} E_i \ , \tag{3.73}$$

alors qu'on trouve en polarisation π:

$$E_t = \frac{2n_1}{n_1 + n_2} E_i \tag{3.74}$$

$$E_r = \frac{n_2 - n_1}{n_1 + n_2} E_i \ , \tag{3.75}$$

Ces deux résultats devraient être identiques, les polarisations π et σ étant physiquement indistinguables. La différence de signe sur l'amplitude réfléchie provient d'une différence dans nos choix d'orientation pour les deux polarisations. Quand l'angle d'incidence tend vers zéro, E_i et E_r ont la même orientation de référence en polarisation σ, alors qu'ils sont opposés en polarisation π. Assez étrangement, ces conventions de signe quelque peu absurdes ont survécu à des générations de manuels...

On peut déduire de là les coefficients de transmission et de réflexion pour l'énergie. Les quantités physiquement intéressantes sont les flux normaux à la surface:

$$n_x \cos\theta_x |E_x|^2 \ , \tag{3.76}$$

avec $x = i, r, t$.Soit R le rapport des flux normaux réfléchi et incident, T le rapport des flux transmis et incident. Leurs expressions sont complexes, mais faciles à obtenir à partir des coefficients de Fresnel. En incidence normale, on trouve:

$$R = \left|\frac{n_1 - n_2}{n_1 + n_2}\right|^2 \tag{3.77}$$

$$T = \frac{4n_1 n_2}{|n_1 + n_2|^2} \ , \tag{3.78}$$

qui vérifient évidemment $R+T = 1$ (cette relation, exprimant la conservation de l'énergie, est d'ailleurs vérifiée quelque soit l'incidence)

Notons qu'en polarisation π, E_r s'annule si:

$$n_2 \cos\theta_i = n_1 \cos\theta_t \ . \tag{3.79}$$

En combinant cette relation avec la troisième loi de Descartes, on voit que cela correspond à

$$\cos\theta_t = \sin\theta_i \tag{3.80}$$

$$\sin\theta_t = \cos\theta_i \tag{3.81}$$

soit encore

$$\theta_t + \theta_i = \pi/2 \tag{3.82}$$

et

$$\tan\theta_i = \frac{n_2}{n_1} \ . \tag{3.83}$$

Pour cette incidence particulière, appelée incidence de Brewster, les rayons réfractés et réfléchis sont perpendiculaires et l'amplitude réfléchie est strictement nulle. On peut d'ailleurs se faire une image assez intuitive du mécanisme de l'incidence de Brewster. Physiquement, l'onde réfléchie doit être rayonnée par des dipôles situés au voisinage de l'interface, dans le matériau 2. Ces dipôles sont orientés selon la polarisation de l'onde réfractée. A l'incidence de Brewster, ils sont donc alignés avec la direction de propagation de l'onde réfléchie. Le rayonnement d'un dipôle étant nul le long de son axe, on comprend bien que l'intensité réfléchie doive être nulle dans ce cas.

On combat au moyen de l'incidence de Brewster les réflexions parasites qui sont sources de pertes dans les systèmes optiques. Les tubes à plasma des lasers à gaz, par exemple, qui contiennent le milieu amplificateur dans la cavité laser, sont fermés par des fenêtres à l'incidence de Brewster, qui n'introduisent donc aucune perte pour la polarisation convenable. La conséquence en est que le rayonnement de ces lasers présente une polarisation linéaire π.

Notons finalement que les relations de Fresnel s'appliquent tout aussi bien au cas des indices de réfraction complexes. Prenons par exemple le cas de l'incidence normale sur un métal à basse fréquence, depuis l'air. L'indice n_1 est alors égal à 1 et l'indice n_2 à:

$$n_2 = (1 + i)x \ , \tag{3.84}$$

avec:

$$x = \sqrt{\frac{\epsilon_r \omega_p^2}{2\gamma\omega}} \gg 1 \ . \tag{3.85}$$

On en déduit le coefficient de réflexion pour l'énergie:

$$R = \frac{(1-x)^2 + x^2}{(1+x)^2 + x^2} \simeq 1 - \frac{2}{x} \ . \tag{3.86}$$

Le coefficient de réflexion est donc très voisin de 1, la différence s'exprimant facilement en fonction des caractéristiques du conducteur ou de la profondeur de peau δ:

$$1 - R = 2k_0\delta \ . \tag{3.87}$$

Plus grande est la conductivité, plus petite est la profondeur de peau et meilleure la qualité du miroir métallique.